THE WARING PROBLEM

by

Wolfgang Joseph Thron

A minor thesis presented to the faculty of the Rice Institute.

July, 26, 1942.

Degree granted 1943.
Most of the material presented in this paper has been taken out of either Landau's "Vorlesungen" (Sections 2 to 6 and 14) or his "Neuere Ergebnisse" (Sections 7 to 12). Other books and some original papers were also consulted.

It has been my main concern to give as complete as possible a proof of Vinogradow's Theorem. To this end Sections 2, 5 and 6 were included. Further most proofs are given in greater detail, than even Landau thought necessary. In other words the paper should now be understandable to Brunk, Cowling and Piranian, even though they will probably never read it.

Other results were only included, if they could be proved easily.
Throughout this paper we shall use the following notation.

Small roman letters, except e and i, denote integers (positive, negative or zero). The letters e and i have their usual meaning, that is $e = \lim(1 + \frac{1}{n})^n$ and $i = \sqrt{-1}$. The letter $p$ always stands for a prime number, where we define as a prime number any positive integer, that is divisible only by $\pm 1$ and by itself.

For $a \neq 0$ the symbol $a / b$, or in words "$a$ divides $b$", is understood to mean: there exists an integer $c$, such that

$$ac=b.$$ 

We permit $c \neq 0$, hence for every $a \neq 0$ $a / 0$.

If $a$ does not divide $b$ we write $a \not| b$.

By $(a,b)$ we denote the greatest common divisor of $a$ and $b$. That is, $d=(a,b) \geq 1$, if $d / a$, $d / b$ and if for every number $f > d$ at least one of the two relations $f / a$, $f / b$ holds.

If $\xi$ is any real number we denote by $[\xi]$ the greatest integer contained in $\xi$. That is

$$[\xi] \leq \xi \leq [\xi] + 1.$$
By $\{\xi\}$ we denote the distance from $\xi$ to the nearest integer. That is

$$\{\xi\} = \min (\lfloor \xi \rfloor + 1 - \xi, \xi - \lfloor \xi \rfloor).$$

Introducing new expressions we may write

$$\lambda = \lambda (k, N),$$

this means $\lambda$ is a function of $k$ and $N$ alone.

Constants like $c_1(k)$, $c_2(k)$, ... are always assumed to be real positive numbers, not necessarily integers.

---

Section 1.

**INTRODUCTION**

In 1636 Fermat announced and claimed to have proved the theorem that any positive integer can be expressed as the sum of four squares of integers (or zero).

In 1770 Waring conjectured that every positive integer is the sum of four squares, nine cubes, 19 fourth powers "and so on". This is Waring's only contribution to the problem that bears his name, for he never proved even a part of his statement. His words do not even outline the problem precisely. To do this we make three definitions.

**Definition 1.1. Let** $g(k)$ **be the smallest integer** $s$

for which the diophantine equation

$$(1.1) \quad n = \sum_{m=1}^{s} x_m^k, \quad k \geq 1,$$
has a solution \( x_1, \ldots, x_s \), with \( x_m \geq 0 \), for every \( n \geq 0 \).

If no such \( s \) exists, we define \( g(k) = \infty \).

**Definition 1.2.** Let \( G(k,m) \) be the smallest integer \( s \) (\( G(k,m) = \infty \), if no \( s \) exists) for which the equation (1.1) has a solution \( x_1, \ldots, x_s, x_j \geq 0 \), for all \( n \geq m \).

We now note, that for \( m_1 > m_2 \)

\[ 1 \leq G(k,m_1) \leq G(k,m_2) \]

if \( G(k,m_2) < \infty \). Hence, \( \lim_{m \to \infty} G(k,m) \) exists, it is equal to a certain integer \( G(k) \). As \( G(k,m) \) takes on only integral values, there exists a certain integer \( n_0(k) \), such that for \( m \geq n_0(k) \) \( G(k,m) = G(k) \). If \( G(k,m_2) = \infty \) we have for all \( m \geq 0 \) \( G(k,m) = \infty \) and define \( G(k) = \infty \).

**Definition 1.3.** Let \( G(k) \) be the smallest integer \( s \) (\( G(K) = \infty \), if no \( s \) exists) for which the equation (1.1) has a solution \( x_1, \ldots, x_s \), with \( x_j \geq 0 \), for all \( n \geq n_0(k) \)

\( n_0(k) \) to be chosen as indicated above, \( n_0(k) = 0 \) for \( G(k) = \infty \).

As Waring's Problem we then define the determination of \( g(k) \) and \( G(k) \) for all \( k \geq 1 \). For \( k=1 \) we have (as Landau does not fail to point out) evidently \( g(1) = G(1) = 1 \). For \( k=2 \) Lagrange proved in 1770 that \( g(2) \leq 4 \) and the number

\[ 7 = 2^2 + 1^2 + 1^2 + 1^2 \]

allows no decomposition into less than four squares, hence \( g(2) = 4 \). For \( k=3 \) Wieferich found \( g(3) = 9 \).
The first general result was Hilbert's proof in 1909 that \( g(k) < \infty \) and hence \( G(k) < \infty \), for all \( k \).

Between 1920 and 1925 Hardy and Littlewood published a series of papers, in which, using a new method developed by them, they gave upper bounds for \( G(k) \). Their best result, obtained in 1925, is, for \( k \geq 4 \)
\[
G(k) < (k-2)2^{k-4} k + 5 + \sum_{\nu} \nu^2 \log 2.
\]
Improving the method of Hardy and Littlewood, Vinogradov obtained in 1935 the excellent result (considering that \( G(k) \geq k+1 \) as we prove in Section 14) that for \( k \geq 4 \)
\[
G(k) < V(k) = \left\lfloor k(6 \log k + 3 \log 6 + 4) \right\rfloor - 2.
\]
It is the proof of approximately this result, that will take up the greatest part of this thesis.

We believe this choice is justified, because Vinogradov's Theorem, evidently includes Hilbert's Theorem, it gives the best upper bound for \( G(k) \), known so far, it is of interest in the theory of diophantine equations, as of course all results on Waring's Problem are, and finally it was Vinogradow's Theorem that enabled Dickson and Pillai independently to determine the exact value of \( g(k) \) for almost all \( k \).

They both proved in 1936, that for \( k > 6 \)
\[
g(k) = l(k) = 2^k q - 2, \quad q = \left\lfloor \left(\frac{3}{2}\right)^k \right\rfloor,
\]
if $2^k \geq q + r + 3$, where $3^k = 2^k q + r$, $0 \leq r < 2^k$. This inequality is satisfied for all $6 < k \leq 400$ and for most greater numbers, possibly for all. For some of the remaining numbers, if any, the value of $g(k)$ has also been determined by Dickson and Pillai. It is thus only for $k = 4, 5, 6$ and possibly a few $k > 400$ that $g(k)$ is still unknown.

These striking results were obtained from Vinogradov's Theorem in the following way. For $k > 6$ Vinogradov's upper bound $V(k)$ is smaller than $I(k)$, which is known to be the lower bound for $g(k)$ (see Section 11), the same was not true even for Hardy's and Littlewood's best result. The problem of finding $g(k)$ is thus reduced to the problem of determining an $N(k)$ as small as possible, such that all $n \leq N(k)$ can be expressed as the sum of $V(k)$ $k$ th powers, and then considering the decomposition of the remaining $n < N(k)$. This is exactly what Dickson and Pillai did.

There have been numerous other investigations in connection with Waring's Problem. $G(4)$ for example was closely investigated by Hardy and Littlewood, and more recently by Heilbronn and Davenport and independently by Esterman, they found $G(4) \leq 17$. Dickson has also made a study of the decomposition of numbers into pairs of $k$ th powers. As a final example we may mention the investigations concerning the decomposition of a number in the form $\sum_{i=1}^{k} f(x_i)$, where $f(x)$ is an arbitrary polynomial with integer coefficients.
Section 2.

CONGRUENCES.

Certain results from the theory of congruences will be needed throughout the remainder of this thesis, these we develop here, without in any way trying to give a complete account of the theory.

Definition 2.1. If \( m > 0 \), we say \( a \) is congruent \( b \) modulo \( m \), and write
\[
a \equiv b \pmod{m},
\]
if and only if \( m \parallel (a-b) \).

Theorem 2.1. The equivalence relation thus defined is reflexive, symmetric and transitive.

Proof. We have for any \( m > 0 \) and any \( a, m \notparallel (a-a) \), hence the relation is reflexive. Further if \( m \parallel (a-b) \) then \( m \parallel (b-a) \), this proves the symmetric property.

The transitivity we prove as follows: \( m \parallel (a-b) \) and \( m \parallel (b-c) \), hence \( m \parallel ((a-b)+(b-c)) \), that is \( m \parallel (a-c) \).

With the help of this equivalence relationship we can divide the integers into classes, which shall be called residue classes mod \( m \). Two numbers \( a \) and \( b \) belong into the same class if and only if they are congruent mod \( m \).

Theorem 2.2. A necessary and sufficient condition for
\[
a \equiv b \pmod{m}
\]
is that
\[
a = mq_1 + r, \quad 0 \leq r < m,
b = mq_2 + r.
\]
Proof. We first note that any number \( a \) can be uniquely represented in the form
\[
a = mq_1 + r , \quad 0 \leq r < m.
\]
If now
\[
b = mq_2 + r ,
\]
then we have
\[
a - b = m(q_1 - q_2)
\]
and hence \( m \mid (a-b) \).
Secondly if \( a = mq_1 + r \) and \( a \equiv b \pmod{m} \), then \( (a-b) = mq \) and
\[
b = a - mq = m(q_1 - q) + r = mq_2 + r .
\]
This completes the proof of the theorem.

It follows from this theorem, that all numbers of one residue class have the same residue \( r \) and that numbers of different classes have different residues. Hence there are \( m \) different residue classes \( \mod m \).

The following theorems are concerned with the manipulations that are possible with congruences.

Theorem 2.3. From
\[
a \equiv b \pmod{m} ; \quad c \equiv d \pmod{m}
\]
follows
\[
a + c \equiv b + c \pmod{m}.
\]

Proof. We have \( m \mid (a-b) \) and \( m \mid (c-d) \), hence
\[
m \mid ((a-b) + (c-d)) \quad \text{that is} \quad m \mid ((a+c) - (b+d)).
\]

Theorem 2.4. From
\[
a \equiv b \pmod{m}
\]
follows for every \( c \)
\[
c \cdot a \equiv b \cdot c \pmod{m}.
\]
Theorem 2.5. From
\[ a \equiv b \pmod{m}; \quad c \equiv d \pmod{m} \]
follows
\[ ac \equiv bc \pmod{m}. \]

Proof. From the preceding theorem it follows that
\[ ac \equiv bc \pmod{m}; \quad bc \equiv bd \pmod{m}. \]
The transitive property of congruences then assures the result.

Theorem 2.6. Let
\[ f(x) = c_0 + c_1 x + \cdots + c_n x^n, \]
and let
\[ a \equiv b \pmod{m}, \]
then
\[ f(a) \equiv f(b) \pmod{m}. \]

Proof. The theorem follows from repeated applications of the preceding theorems.

This theorem justifies the following definition.

Definition 2.2. As the number of solutions of the congruence
\[ f(x) \equiv 0 \pmod{m} \]
we regard the number of solutions \( x_0 \), where \( 0 \leq x_0 < m. \)
This is the same as the number of residue classes \( x \).
mod m all of whose elements satisfy the congruence
(by the preceding theorem either all or no elements of
a certain residue class satisfy a given congruence).

Theorem 2.7. If \((c, m) = 1\), and
\[ ac \equiv bc \pmod{m}, \]
then
\[ a \equiv b \pmod{m}. \]

Proof. We have \(m \mid ac - bc\), and hence \(m \mid (a-b)c\),
as \((c, m) = 1\) the preceding relation implies \(m \mid (a-b)\).

Theorem 2.8. From
\[ ac \equiv bc \pmod{m} \]
follows
\[ a \equiv b \pmod{\left(\frac{m}{(c, m)}\right)}. \]

Proof. \(m \mid (a-b)c\) hence \(\frac{m}{(c, m)} \mid (a-b)\frac{c}{(c, m)}\).
Now \(\left(\frac{m}{(c, m)}, \frac{c}{(c, m)}\right) = 1\) and hence \(\frac{m}{(c, m)} \mid (a-b)\).

Theorem 2.9. If \(n \mid m\) and \(ac \equiv bc \pmod{m}\),
then
\[ a \equiv b \pmod{m}. \]

Proof. \(m \mid (a-b)\) and \(n \mid m\) hence \(n \mid (a-b)\).

Theorem 2.10. If
\[ a \equiv b \pmod{m}, \]
then
\[ ac \equiv bc \pmod{m^2}, \]
and conversely.

Proof. From \(m \mid (a-b)\) follows \(mc \mid (a-b)c\),
\(mc \mid (ac-bc)\) and conversely.
Definition 2.3. The number $a$ is said to be prime to $b$ if and only if $(a,b)=1$.

Definition 2.4. The function $\varphi(m)$ is equal to the number of integers $0 \leq a < m$, which are prime to $m$.

Theorem 2.11. $\varphi(p) = p-1$.

Proof. No number greater than zero and smaller than $p$ is divisible by $p$.

Theorem 2.12. $\varphi(p^h) = p^h - p^{h-1}$.

Proof. Among the numbers $0 < a < p^h$ there are exactly $p^{h-1}$ numbers divisible by $p$. These are the numbers of the form $pb$ where $b$ runs from 1 to $p^{h-1}$. Hence $(p^h) = p^h - l - (p^{h-1}-l) = p^h - p^{h-1}$.

Theorem 2.13. If
\[ a \equiv b \pmod{m}, \]
then
\[ (a,m) = (b,m). \]

Proof. We have $b = a + mq$, hence $(a,m) / b$. Evidently $(a,m) / m$ and hence $(a,m) / (b,m)$. Similarly it can be shown that $(b,m) / (a,m)$. It follows that $(b,m) = (\varphi,m)$.

This theorem allows us to speak of residue classes mod $m$ as being prime to $m$. For if one element of a class satisfies $(a,m)=1$, then this condition is also satisfied by all the others. And conversely if no one element of a class is not prime to $m$, then none of the elements is prime to $m$. 
Definition 2.5. A complete residue system mod m is a set of m integers, no two of which are congruent mod m. This is the same as saying, that every residue class is represented by exactly one element of the set.

Definition 2.6. A reduced residue system mod m is a set of \( \varphi(m) \) integers, all of which are prime to m, and no two of which are congruent mod m.

Theorem 2.14 If \((k, m) = 1\) and if the numbers \(a_1, a_2, \ldots, a_m\) form a complete residue system mod m, then the m numbers \(a_1k, a_2k, \ldots, a_mk\) form a complete residue system mod m.

Proof. From Theorem 2.7 it follows, that

\[ a_{rk} \equiv a_{sk} \pmod{m} \]

implies

\[ a_r \equiv a_s \pmod{m}, \]

that means \(a_r = a_s\). Hence no two of the m numbers \(a_{sk}\) are congruent mod m. The theorem then follows from Definition 2.5.

Theorem 2.15. If \(y\) runs over a complete residue system mod \(q_1\) and \(z\) runs over a complete residue system mod \(q_2\), where \((q_1, q_2) = 1\), then \(zq_1 + yq_2\) runs over a complete residue system mod \(q_1q_2\).

Proof. The expression \(zq_1 + yq_2\) takes on \(q_1q_2\) different values. It is sufficient to show that no two
of them are congruent mod q_1q_2. Assume
\[ zq_1 + yq_2 \equiv z'q_1 + y'q_2 \pmod{q_1q_2} \]
then \( q_1q_2 / ((z-z')q_1 (y-y')q_2) \). Recalling that
\((q_1, q_2) = 1 \) we must have \( q_1 / (y-y') \) and \( q_2 / (z-z') \),
this is equivalent to
\[ y \equiv y' \pmod{q_1}; \quad z \equiv z' \pmod{q_2}. \]
Thus \( y=y' \) and \( z=z' \).

**Theorem 2.16.** Let \( d \mid m \), then a residue class
\[ \text{mod } \frac{m}{d} \]
is composed of exactly \( d \) residue classes \( \text{mod } m \).

**Proof.** A residue class \( \text{mod } \frac{m}{d} \) consists of all
numbers \( \frac{m}{d} + r \), where \( r \) is fixed \((0 \leq r < \frac{m}{d})\), and \( q \)
takes on all integral values. We can write
\[ q = gd + s, \quad s = 0, 1, \ldots, d-1, \]
and hence
\[ \frac{m}{d} + r = gm + s + r \equiv gm + r'. \]
Here \( r' \) can take on \( d \) different values and we will always
have \( 0 \leq r' < m-1 \).

**Theorem 2.17.** If \( (k, m) = 1 \) and if \( a_1, a_2, \ldots, a_{\varphi(m)} \)
form a reduced residue system \( \text{mod } m \), then \( a_1k, \ldots, a_{\varphi(m)}k \)
form a reduced residue system \( \text{mod } m \).

**Proof.** From \( (k, m) = 1 \) and \( (a_r, m) = 1 \) follows
\( (a_rk, m) = 1 \). Further according to Theorem 2.14 no two
numbers of the new set are congruent \( \text{mod } m \). The theorem
then follows from Definition 2.6.
Theorem 2.18. If $y$ runs over a reduced residue system mod $q_1$ and $z$ runs over a reduced residue system mod $q_2$, where $(q_1,q_2)=1$, then $zq_1 + yq_2$ runs over a reduced residue system mod $q_1q_2$. It follows that for $(q_1,q_2)=1$

$$\varphi(q_1) \cdot \varphi(q_2) = \varphi(q_1q_2).$$

Proof. It follows from Theorem 2.15 that as $y$ and $z$ run over complete residue systems so does $zq_1 + yq_2$. It is sufficient to show, first that if either $(z,q_2)>1$ or $(y,q_1)>1$, then $(zq_1 + yq_2, q_1q_2)>1$, this is obvious. Secondly if both $(z,q_2)=1$ and $(y,q_1)=1$, then $(zq_1 + yq_2, q_1q_2)=1$, this too is evident if we recall, that $(q_1,q_2)=1$.

Theorem 2.19. If $(a,m)=1$ the congruence

$$ax + a_0 \equiv 0 \pmod{m}$$

has exactly one solution.

Proof. According to Theorem 2.14 the set $a_0, a_1, ..., a(m-1)$ forms a complete residue system mod $m$, exactly one of these numbers therefore is $\equiv -a_0 \pmod{m}$.

Theorem 2.20 The congruence

$$(2.1) \quad ax + a_0 \equiv 0 \pmod{m}$$

has a solution if and only if $(a,m) / a_0$. The number of solutions is then $(a,m)$, and all $x$ of a certain residue
class mod \( \frac{m}{(a,m)} \) are solutions of the congruence.

**Proof.** If (2.1) is solvable, we have according to Theorem 2.9

\[ ax + a_0 \equiv 0 \pmod{(a,m)}, \]

now

\[ ax \equiv 0 \pmod{(a,m)} \]

and hence

\[ a_0 \equiv 0 \pmod{(a,m)}. \]

This proves the first part of the theorem. If then \( a_0 \equiv 0 \pmod{(a,m)} \) the congruence

\[ \frac{a}{(a,m)}x + \frac{a_0}{(a,m)} \equiv 0 \pmod{m} \]

has according to Theorem 2.19 exactly one solution.

For all elements of this residue class \( \text{mod} \frac{m}{(a,m)} \) the congruence (2.1) is satisfied, by Theorem 2.10. This residue class \( \text{mod} \frac{m}{(a,m)} \) consists by Theorem 2.16 of exactly \( (a,m) \) residue classes \( \text{mod} m \). According to Definition 2.2 the number of solutions of the congruence is therefore at least equal to \( (a,m) \).

That there are no other solutions, follows from the fact that not only does (2.2) imply (2.1), but also conversely.

**Theorem 2.21. (Fermat)** If \( (a,m)=1 \), then

\[ a^{\phi(m)} \equiv 1 \pmod{m}. \]

**Proof.** Let \( a_1, a_2, \ldots, a^{\phi(m)} \) be a reduced residue system \( \text{mod} m \). According to Theorem 2.17 the set \( a_1a_1, a_2, \ldots, a^{\phi(m)} \) also forms a reduced residue
system mod m. Every element of one set is then congruent
to one element of the other set and from Theorem 2.5
it then follows that
\[ \varphi(m) \bigg| \prod_{k=1}^{n} a_k \equiv \prod_{k=1}^{n} a_k \pmod{m}, \]
This can be written as
\[ \varphi(m) \bigg| a_k \equiv \varphi(m) a_k \pmod{m}. \]
Now \( \prod_{k=1}^{n} a_k, m \equiv 1 \), hence Theorem 2.7 can be applied to
yield the announced result.

**Theorem 2.22.** If \( p \not| a \), then
\[ a^{p-1} \equiv 1 \pmod{p}. \]

**Proof.** \( p \not| a \) implies \( (a, p) = 1 \). By Theorem 2.11
\( \varphi(p) = p - 1 \), hence this theorem follows from the preceding one.

**Definition 2.7.** Let \( (a, m) = 1 \). The number \( a \)
is said to belong to the exponent \( f \) mod \( m \), if among the
powers \( a^1, a^2, \ldots, a^f \) is the first power, such that
\[ a^f \equiv 1 \pmod{m}. \]

It follows from Fermat's Theorem, that an \( f \)
always exists.

**Theorem 2.23.** If \( a \) belongs to the exponent \( f \)
mod \( m \), then for \( b_1 \geq 0, b_2 \geq 0 \)
\[ a^{b_1} \equiv a^{b_2} \pmod{m} \]
if and only if
\[ b_1 \equiv b_2 \pmod{f}. \]
Further for \( b \geq 0 \)
\[ a^b \equiv 1 \pmod{m} \]
if and only if $f / b$. Hence by Fermat's Theorem $f / \varphi(m)$.

Proof. Let $b_2 > b_1$, then as $(a,m)=1$ Theorem 2.7 leads to

$$a^{b_2-b_1} \equiv 1 \pmod{m}.$$

Now let $b_2 - b_1 = qf + r$, $0 \leq r < f$. Then

$$1 \equiv a^{b_2-b_1} \equiv a^{qf}a^r \equiv a^r \pmod{m}.$$

Hence from the definition of $f$ it follows that $r=0$ and $f / b_2 - b_1$. On the other hand if $b_2 - b_1 = qf$ we have

$$a^{b_2} \equiv a^{b_1+qf} \equiv a^{b_1} \pmod{m}.$$

The second part of the theorem is obtained by setting $b=b_1$ and $0=b_2$.

Theorem 2.24. Let $p \not| \, c_n$ and let

$$f(x) = c_0 + c_1 x + \ldots + c_n x^n,$$

then the congruence

$$f(x) \equiv 0 \pmod{p}$$

has at most $n$ solutions.

Proof. The theorem is true for $n=0$. Assume it is true for all $f(x)$ of degree $n-1$. We then show, by contradiction, that it is also true for all $f(x)$ of degree $n$. Assume there is an $f(x)$ of degree $n$ such that for $n+1$ numbers $x_1, x_2, \ldots, x_{n+1}$ the congruence is satisfied. Then

$$f(x) - f(x_1) = (x - x_1)g(x),$$

where $g(x) = b_0 + \ldots + c_n x^{n-1}$. Now for $x_2$ to $x_{n+1}$
\[ f(x_k) - f(x_1) \equiv (x_k - x_1)g(x_k) \equiv 0 \pmod{p} \]

but \((x_k - x_1) \neq 0 \pmod{p}\), otherwise they would not be different solutions, in the sense in which we use the word, hence

\[ g(x) \equiv 0 \pmod{p} \]

for \(n\) different values, this however contradicts our assumption. (The argument, that if a product is congruent to 0, then one of the factors is congruent to 0, holds of course only for congruences \(\text{mod } p\), where \(p\) is a prime.)

**Theorem 2.25.** Let \(q\) be a prime number, let \(h > 0\) be such that \(q^h / p - 1\). Then there exists an \(a\) that belongs to \(q^h \pmod{p}\).

**Proof.** According to the preceding theorem the congruence

\[ x^{\frac{p-1}{q}} \equiv 1 \pmod{p} \]

has at most \(\frac{p-1}{q} < p - 2\) solutions. Hence there exists at least one \(c\) such that \((c \neq 0 \pmod{p})\)

\[ \exists \quad c^{\frac{p-1}{q}} \not\equiv 1 \pmod{p}. \]

Let

\[ a = c^{\frac{p-1}{q}}. \]

then

\[ a^h \equiv c^{p-1} \equiv 1 \pmod{p}. \]

Let \(a\) belong to \(f\), then by Theorem 2.23 \(f / q^h\). If \(f \neq q^h\) then \(f / q^{h-1}\). This however would lead to
\[ a^{\frac{p-1}{h-1}} \equiv c^q \equiv 1 \pmod{p}, \]

this is impossible, and therefore is the desired number, because it belongs to \( q^h \).

**Theorem 2.26.** There exists a \( g \) that belongs to \( p-1 \pmod{p} \).

**Proof.** If \( p=2 \) we can set \( g=1 \). If \( p>2 \) we have

\[ p-1 = \prod_{n=1}^{m} p_n^{h_n} \quad (p_m \neq p_1) \]

For \( r=1 \) Theorem 2.25 gives the desired number. If \( r>1 \) we choose according to Theorem 2.25 for every \( n=1,2,\ldots, r \) an \( a_n \) which belongs to \( p_n^{h_n} \). We then set

\[ g = \prod_{n=1}^{r} a_n \]

and call the exponent to which \( g \) belongs \( f \). According to Theorem 2.23 \( f \mid p-1 \). If \( f \neq p-1 \), then \( f \mid \frac{p-1}{p_1} \) (that is \( f \mid \frac{p-1}{p_m} \) and we set \( m=1 \)). Now \( p_n^{h_n} \mid \frac{p-1}{p_1} \) for \( n=2,3,\ldots, r \) and hence

\[ 1 \equiv \frac{p-1}{p_1} \equiv a_1 \frac{p-1}{p_1} \prod_{n=2}^{r} a_n \frac{p-1}{p_1} \equiv a_1 \frac{p-1}{p_1}. \pmod{p} \]

Hence as \( a_1 \) belongs to \( p_1^{h_1} \), we have by Theorem 2.23 \( p_1^{h_1} \mid \frac{p-1}{p_1} \). This is a contradiction and hence \( f=p-1 \), and \( g \) belongs to \( p-1 \).

**Definition 2.8.** Every number \( g \) that belongs to \( p-1 \pmod{p} \) is called a primitive root modulo \( p \).
The previous theorem insures the existence of at least one primitive root for every \( p \).

**Theorem 2.27.** If \( p > 2, h > 0 \), there exists a number \( g \) that belongs to \( \varphi(p^h) \) mod \( p^h \).

**Proof.** The case \( h=1 \) is taken care of by the preceding theorem. \( (\varphi(p)=p-1) \). Let \( h > 1 \) and let \( g \) be a primitive root mod \( p \). One can choose \( g \) in such a way that

\[
g^{p-1} \not\equiv 1 \pmod{p^2}.
\]

This follows from the following consideration: if \( g \) is a primitive root mod \( p \) so is \( g+p \). And if

\[
g^{p-1} \equiv 1 \pmod{p^2}
\]

then

\[
(g+p)^{p-1} \equiv g^{p-1}(p-1)g^{p-2}p \equiv 1 + (p-1)g^{p-2}p \pmod{p^2},
\]

and as \( p \nmid p-1, p \nmid g \), it follows that

\[
(g+p)^{p-1} \not\equiv 1 \pmod{p^2}.
\]

We now show, that the number \( g \) given by (2.3) is the desired number.

We first show, that for all \( h > 1 \)

\[
g^{p^{h-2}}(p-1) \not\equiv 1 + k_{h}^{p} \pmod{p^h}; p \nmid k_{h}.
\]

For \( h=2 \) Fermat's Theorem gives

\[
g^{p-1} = 1 + k_{2}p,
\]

and relation (2.3) insures that \( p \nmid k_{2} \). Now assume
That (2.4) is true for \( h \), for \( h+1 \) we then have

\[
g_p^{p^{h-1}(p-1)} = (1 + k_h p^{h-1})^p
\]

\[
= 1 + k_h p^h + k_h \frac{p}{2} \frac{p-1}{2} p^2(h-1) + n_p \frac{3}{2}(h-1).
\]

Of the three last terms all except \( k_h p^h \) are divisible by \( p^{h+1} \), if we write these terms now as one term \( k_{h+1} p^h \), then \( p \nmid k_{h+1} \) as \( p \nmid k_h \). Hence relation (2.4) is true for \( h+1 \), if it is true for \( h \). By complete induction it then is true for all \( h \geq 2 \).

Now let \( f \) be the exponent to which \( g \) belongs mod \( p^h \). From Theorem 2.23 it then follows that \( f / p^{h-1}(p-1) \).

Further

\[
g^f \equiv 1 \pmod{p^h},
\]

Hence by Theorem 2.9

\[
g^f \equiv 1 \pmod{p},
\]

and hence as \( g \) belongs to \( p - 1 \) mod \( p \) we have, again from Theorem 2.23, \( p-1 / f \). This implies either

\( f = p^{h-1}(p-1) \) or \( f / p^{h-2}(p-1) \). In the second case we would obtain

\[
g^{p^{h-2}(p-1)} \equiv 1 \pmod{p^h}.
\]

This however contradicts relation (2.4). Therefore

\( f = p^{h-1}(p-1) = \varphi(p^h) \) and \( g \) belongs to \( \varphi(p^h) \) mod \( p^h \).

**Theorem 2.28.** Let \( p > 2, h > 0 \) and let \( g \) belong to \( \varphi(p^h) \) mod \( p^h \), then the set \( g^1, g^2, \ldots, g^{\varphi(p^h)} \)

is a reduced residue system mod \( p^h \).
Proof. According to Definition 2.6 it is sufficient to show, that no two different powers of \( g \) are congruent \( \mod p^h \). This follows from the preceding theorem.

**Theorem 2.29.** If \( m \geq 2 \), then \( 5 \) belongs to \( 2^{m-2} \) \( \mod 2^m \).

**Proof.** For \( m=2 \) the theorem is trivial. For \( m>2 \) we first prove by induction that

\[
5^{2^{m-3}} = 1 + h_m2^{m-1} \mod 2^m.
\]

This is true for \( m=3 \) (\( 5=1 \ 1 \ 4 \)). Now assume (2.5) is true for \( m \), then it is true for \( m+1 \), this is seen as follows

\[
5^{2^m} = (1 + h_m2^{m-1})^2 = 1 + h_m2^m + h_m2^{2m-2} = 1 + h_{m+1}2^m \mod 2^{m+1}.
\]

From (2.5) follows

\[
5^{2^{m-3}} \equiv 1 \mod 2^m
\]

\[
5^{2^{m-2}} \equiv 1 \mod 2^m.
\]

Let \( f \) be the exponent to which \( 5 \) belongs \( \mod 2^m \). From the two preceding congruences follows \( f \not\equiv 2^{m-3}, f \not\equiv 2^{m-2} \), hence \( f = 2^{m-2} \).

**Theorem 2.30.** The set \( 5^1, 5^2, \ldots, 5^{2^{m-2}} \) takes \( \mod 2^m \) all values \( a \equiv 1 \mod 4 \).

**Proof.** All numbers \( 5^1, 5^2, \ldots, 5^{2^{m-2}} \) are incongruent \( \mod 2^m \). This follows from the preceding theorem. All these
numbers are congruent 1 mod 4 \( (5 \equiv 1 \pmod{4}) \) hence
\[ 5^b \equiv 1^b \equiv 1 \pmod{4} \]. Now there are exactly \( 2^{m-2} \) residue classes mod \( 2^m \) that are \( \equiv 1 \pmod{4} \) (Theorem 2.16). That proves the theorem.

It is worth noting, that there is no number \( a \), that belongs to \( 2^{m-1} = \varphi(d) = \varphi(2^m) \mod 2^m \), for \( m > 2 \). This is evident for all \( a \equiv 1 \pmod{4} \) from the preceding theorem. For \( a \equiv -1 \pmod{4} \), we have
\[ a^2 \equiv 1 \pmod{8} \]
hence
\[ a^2 \equiv 5^d \pmod{8} \],
the exponent \( d \) here is even, as \( 5^k \) is congruent 1 mod 8 only for even \( k \). The result then follows easily.

**Theorem 2.31.** The number of solutions of the congruence
\[(2.6) \quad x^k \equiv 1 \pmod{p} \]
is \( (k,p-1) \).

**Proof.** Let \( g \) be a primitive root mod \( p \) and let \( d = (k,p-1) \), then if \( n=1, 2, \ldots, d \) the numbers
\[ g^{n\left(\frac{p-1}{d}\right)} \]
are roots (different because of Theorem 2.28) of the congruence \((2.6)\). We have
\[ 1 \equiv \left(\frac{g^k}{d}\right)^{p-1} \equiv \left(\frac{g^k}{d}\right)^{p-1} \pmod{p}. \]
There are therefore at least \( d \) roots.
Now let $a = g^b$ be a root of (2.6) and let $a$ belong to $f$. From Theorem 2.23 we then have $f / k$ and $f / p-1$, hence $f / d$. Further we evidently have $bf = p-1$ and from $f / d$ we deduce $f = \frac{d}{m}$, where $1 \leq m \leq d$. Hence

$$b = \frac{p-1}{d} m,$$

and $a$ therefore is one of the roots already found. Hence there are exactly $d$ roots.

Definition 2.9. If a residue class mod $m$ has the property that all its members satisfy the congruence

$$x \equiv v^k \pmod{m}$$

for some $v$, the class is called a $k$th power residue class mod $m$.

It is clear that if

$$a \equiv v^k \pmod{m}$$

and

$$a \equiv b \pmod{m},$$

then

$$b \equiv v^k \pmod{m}.$$

That is either all or no element of a class are $k$th power residues, and the same $v$ can be used for all elements of one class.

Section 3.

The Theorem of Lagrange.

We are now in a position to prove the first result, that was obtained in connection with Warings Problem, this is the determination of $g(2)$. 
Theorem 3.1 (Euler's Identity)

\[(3.1) \quad (x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2)\]

\[= \left(x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4\right)^2 \]

\[+ \left(x_1 y_2 - x_2 y_1 + x_3 y_4 - x_4 y_3\right)^2 \]

\[+ \left(x_1 y_3 - x_3 y_1 + x_4 y_2 - x_2 y_4\right)^2 \]

\[+ \left(x_1 y_4 - x_4 y_1 + x_2 y_3 - x_3 y_2\right)^2.\]

Proof. The identity is easily verified by computation.

Theorem 3.2. For every \( p > 2 \) there exists an \( m \) such that \( 1 < m < p \), and

\[mp = x_1^2 + x_2^2 + x_3^2 + x_4^2\]

is solvable.

Proof. The \( \frac{p+1}{2} \) numbers \( x^2, 0 \leq x \leq \frac{p-1}{2} \)

are incongruent mod \( p \). For

\[x_1^2 \equiv x_2^2 \pmod{p}\]

implies \( p \mid (x_1 - x_2)(x_1 + x_2) \) and hence

\[x_1 \equiv \pm x_2 \pmod{p}.\]

The same is true for the \( \frac{p+1}{2} \) numbers \( -1 - y^2, 0 \leq y \leq \frac{p-1}{2} \).

Hence there exists a pair \( x, y \) such that

\[x^2 \equiv -1 - y^2 \pmod{p} \quad |x| < \frac{p}{2}, \quad |y| < \frac{p}{2}.\]

(There are \( p+1 \) numbers \( x^2, -1 - y^2 \), but only \( p \) different residue classes mod \( p \)).

We thus have

\[x^2 + y^2 + 1 + 0^2 = mp\]

\[0 < mp < \frac{p^2}{4} + \frac{p}{4} + 1 < \frac{p^2}{2} + 1 < p^2.\]
Theorem 3.3. For every $p$

\[ p = x_1^2 + x_2^2 + x_3^2 + x_4^2 \]

is solvable.

**Proof.** For $p=2$ we have

\[ 2 = 2^2 = 0^2 + 0^2. \]

Now let $p > 2$. Let $m = m(p)$ be the smallest number such that

\[ (3.2) \quad mp = x_1^2 + x_2^2 + x_3^2 + x_4^2 \]

is solvable. According to the preceding theorem $m < p$.

We want to show, that $m = 1$. We first prove that $m$ is odd.

Assume $m$ is even, then one can easily see that

\[ x_1 + x_2 + x_3 + x_4 \equiv 0 \pmod{2}, \]

and if the $x_j$ are suitably numbered

\[ x_1 + x_2 \equiv 0 \pmod{2}, \]
\[ x_3 + x_4 \equiv 0 \pmod{2}. \]

It is then easily verified that

\[ \frac{m}{2^p} = \left( \frac{x_1 + x_2}{2} \right)^2 + \left( \frac{x_3 - x_4}{2} \right)^2 + \left( \frac{x_3 + x_4}{2} \right)^2 \]

where the four squares on the right are integers. This is a contradiction, as $m$ was supposed to be the smallest integer having the given property.

We now assume $m > 1$, $m$ odd, hence $m \geq 3$. We choose $y_k$, $k = 1, 2, 3, 4$ in such a way that

\[ y_k \equiv x_k \pmod{m} \quad |y_k| < \frac{m}{2}. \]

This is possible as the numbers between $\frac{m-1}{2}$ and $\frac{m+1}{2}$, including the endpoints, form a complete residue system mod $m$. 
Thus we obtain
\[ \sum_{k} y_k^2 \equiv \sum_{k} x_k^2 \equiv mp \equiv 0 \pmod{m}, \]
(3.3) \[ \sum_{k} y_k^2 = mn < 4 \frac{m^2}{4} = m^2. \]
Here \( m > n > 0 \), otherwise \( y_k = 0 \) for all \( k \) and hence \( m/x_k \) for every \( k \) and hence \( m^2/p \), \( m/p \), which is impossible as \( 1 < m < p \). (This is the only place where \( m \neq 1 \) comes in.)

Now we multiply (3.2) and (3.3), and apply (3.1) to the product, this gives
(3.4) \[ m^2np = \text{right side of (3.1)}, \]
every bracket on the right hand side of (3.4) is divisible by \( m \). For the first we have
\[ \sum_{k} x_k y_k \equiv \sum_{k} x_k^2 \equiv 0 \pmod{m} \]
and for the remaining ones we note that
\[ x_k y_h - x_h y_k \equiv x_k x_h - x_h x_k \equiv 0 \pmod{m}. \]
Dividing (3.4) by \( m^2 \) we get
\[ np = z_1^2 + z_2^2 + z_3^2 + z_4^2 \]
where \( 0 < n < m \), contradicting the minimal property of \( m \).
Hence \( m = 1 \) and the theorem is proved.

Theorem 3.4. \( g(2) = 4 \).

Proof. The theorem is obvious for \( n = 0,1 \). For \( n = 2 \), we note that if the theorem is correct for \( n_1 n \) and \( n_2 \), then by Theorem 3.1 it also is correct for \( n_1 n_2 \). From Theorem 3.3 it follows that the theorem is correct for all \( p \). Every \( n = p_1 \ldots p_r \), the theorem thus follows.
Section 4.

AN OUTLINE OF HILBERT'S PROOF.

In 1909 Hilbert proved, that \( g(k) \) is finite for every \( k \).

Starting from an integral, he was able to derive the following formula, an identity in \( y_j \):

\[
N(y_1^2 + \cdots + y_4^2)^k = \sum_{m=1}^{\infty} (d_m, 1 y_1^2 + \cdots + d_m, 4 y_4^2)^{2k},
\]

where \( N=N(k) > 0 \), \( a=a(k) > 0 \) and \( d_m, j = d_m, j(k) \) are all integers. Using Lagrange's Theorem the finitness of \( g(k) \) for \( k=2^m \) can then be proved by induction. It is true for \( m=1 \), now assume it has been shown that every number \( n \) is decomposable into the sum of \( f(2^m) \) \( 2^m \) th powers. For any multiple of \( N(2^m) \) we then have

\[
nN = \sum_{j=1}^{f} N x_j 2^m.
\]

By Lagrange's Theorem

\[
x_j = y_1, j^2 + \cdots + y_4, j^2,
\]

and hence by formula (4.1)

\[
nN = \sum_{j=1}^{f} \sum_{h=1}^{\infty} z_{j, h} 2^{m+1},
\]

that is \( nN \) is the sum of \( mf \) \( 2^{m+1} \) st powers. Any number \( n = Nq + r, r < N \), and hence is expressible as the sum of \( mf + N - 1 \) \( 2^{m-1} \) st powers.
As a concrete illustration of this method we prove:

**Theorem 4.1.** \( g(4) \leq 53. \)

**Proof.** It can be verified, that for \( k=2 \) the identity (4.1) has the form

\[
6(y_1^2 + y_2^2 + y_3^2 + y_4^2)^2
= (y_1 + y_2)^4 + (y_1 - y_2)^4 + (y_1 + y_3)^4 + (y_1 - y_3)^4 + (y_1 + y_4)^4 + (y_1 - y_4)^4
+ (y_2 + y_3)^4 + (y_2 - y_3)^4 + (y_2 + y_4)^4 + (y_2 - y_4)^4 + (y_3 + y_4)^4 + (y_3 - y_4)^4.
\]

Now every \( m \geq 0 \) can be decomposed into the sum of 4 squares, thus for every \( m \)

\[
6m^2 = h_1^4 + \ldots + h_4^4.
\]

Again, every \( v \) can be expressed as the sum of 4 squares

\[
v = m_1^2 + \ldots + m_4^2,
\]

and hence

\[
6v = 6m_1^2 + \ldots + 6m_4^2 = h_1^4 + \ldots + h_4^4.
\]

Finally every \( n = 6v + r \), where \( 0 \leq r \leq 5 \). \( r \) is the sum of 5 4th powers \((1^4, 0^4)\), and the theorem follows.

For the general case Hilbert developed from (4.1) another formula

\[
(4.2) \quad u = \sum_{h=1}^{R} \frac{1}{S} a_h(y_h^k + z_h^k).
\]

This relation holds for all

\[
(4.3) \quad k(x^k(x^1)^k) \leq u \leq K(x^k + x^1)^k + x^k.
\]

The numbers \( R = R(k) \), \( S = S(k) \), \( K = K(k) \), \( a_h = a_h(k) \)

are integers \((R \geq S, K \text{ are positive})\).
We now note, that from some number $x_0$ on the intervals $J_x$, defined by $(4.3)$, overlap, hence for all $u \geq u_0$, the number $Su$ is expressible as the sum of $2R(k)$ $k$th powers. It then follows, that every number $n \geq 0$ is expressible as the sum of

$$\max(2R(k)+S(k)-1, S(k)u_0(k)-1)$$

$k$th powers.

We give this outline of Hilbert's proof, leaving out all the real difficulties that appear in establishing the relations $(4.1)$ and $(4.2)$, only to contrast it with the, as will subsequently become clear, completely different approach of Hardy and Littlewood and of Vinogradov.

Section 5.

CHARACTERS.

This and the next five sections serve to lay the ground work for the proof of Vinogradov's Theorem.

**Definition 5.1.** A function with integral arguments and complex values is called a character, denoted by $\chi(a)$, mod $k$, if it satisfies the following four conditions:

1) $\chi(a) = 0$, if $(a,k) > 1$;
2) $\chi(1) \neq 0$;
3) $\chi(a_1a_2) = \chi(a_1)\chi(a_2)$;
4) $\chi(a_1) = \chi(a_2)$, if $a_1 \equiv a_2 \pmod{k}$. 


Theorem 5.1. For every character \( \chi(1) = 1 \).

Proof. According to iii) \( \chi(1) = \chi(1 \cdot 1) = \chi(1) \cdot \chi(1) \), and from ii) it follows that \( \chi(1) \neq 0 \), hence \( \chi(1) = 1 \).

Theorem 5.2. If \( (a, k) = 1 \), then \( (\chi(a))^{\varphi(k)} = 1 \), \( \chi(a) \) is a \( \varphi(k) \) th unit root and \( \chi(a) = 1 \).

Proof. Fermat's Theorem gives \( a^{\varphi(k)} = 1 \pmod{k} \).

From iii) it follows that
\[ (\chi(a))^{\varphi(k)} = \chi(a^{\varphi(k)}) \]
Condition iv) combined with Theorem 5.1 then yields
\[ \chi(a^{\varphi(k)}) = \chi(1) = 1 \]
This proves the theorem.

Theorem 5.3. For every \( k \) there exists only a finite number of different characters \( \pmod{k} \); there exists however at least one.

(Two functions \( \chi_1(a) \) and \( \chi_2(a) \) are called different, if there exists at least one \( a \) for which \( \chi_1(a) \neq \chi_2(a) \).)

Proof. According to i) and Theorem 5.2 there exist only a finite number of values of \( \chi(a) \) for every \( a \). From iv) it follows that \( \chi(a) \) is completely determined, if its values for \( 1 \leq a \leq k \) are given. This proves, that there is only a finite number of different functions \( \chi(a) \). The function
(5.1) \[ \chi_0(a) = \begin{cases} 0 & \text{for } (a,k) > 1 \\ 1 & \text{for } (a,k) = 1 \end{cases} \]

is a character, it satisfies all the conditions of Definition 5.1. This completes the proof of the theorem.

**Definition 5.2.** The character \( \chi_0(a) \) is called the main character.

**Theorem 5.4.** If \( \chi(a) \) is a character, so is \( \overline{\chi}(a) \).

**Proof.** The four conditions i) to iv) are evidently fullfilled for \( \overline{\chi}(a) \), if they are fulfilled for \( \chi(a) \).

**Theorem 5.5.** If \( a \) runs over a complete residue system mod \( k \), then

\[ \sum_a \chi(a) = \left\{ \begin{array}{ll} \varphi(k) & \text{for } \chi_0 \\ 0 & \text{otherwise.} \end{array} \right. \]

**Proof.** The function \( \chi_0(a) \) is equal to zero for all \( a \) except the \( \varphi(k) \) numbers \( a \), with \( (a,k) = 1 \). In those cases \( \chi_0(a) = 1 \). Hence

\[ \sum_a \chi_0(a) = \varphi(k). \]

If there exists any other characters, then there must exist, for each of these characters, a number \( b \), with \( (b,k) = 1 \), such that \( \chi(b) \neq 1 \). According to Theorem 2.14 the numbers \( ba \) then run over a complete residue system mod \( k \), if the numbers a dom. Hence from condition iv) it follows that
\[ \xi = \sum_{a} \chi(\frac{\xi}{a}) = \sum_{a} \chi(ba) \]

Condition iii) then gives

\[ \xi = \sum_{a} \chi(ba) = \sum_{a} \chi(b) \chi(a) = \chi(b) \sum_{a} \chi(a) = \chi(b) \xi. \]

This leads to

\[ (\chi(b) - 1) \xi = 0. \]

And as \( \chi(b) - 1 \neq 0 \) by definition, it follows that \( \xi = 0. \)

Section 6.

**Farey Series.**

We present here another device, that proved helpful in the solution of Waring's Problem. Hardy and Littlewood were the first to use it in that connection.

**Definition 6.1.** Let \( \mathbb{Q} \) be fixed, the series

formed by all the fractions

\[ \frac{a}{b}, (a,b)=1, 0 \leq a \leq b \leq n, b > 0, \]

is called the Farey Series Belonging to \( n \), if the fractions are ordered according to size.

**Theorem 6.1.** Let \( \frac{a}{b} < \frac{a'}{b'} \) be two consecutive members of the Farey Series belonging to \( n \), then

\[ b + b' \geq n + 1 \]

and

\[ ba' - ab' = 1. \]

**Proof.** We first show that there exists a number pair \( x, y \), with \( n - b < y \leq n \), such that
(6.1) \[ bx - ay = 1. \]

According to Theorem 2.19 the congruence

(6.2) \[ ay + 1 \equiv 0 \pmod{b} \]

has one solution \( y \), as \( (a, b) = 1 \). With \( y \) any number congruent \( y \mod b \) is also a solution of this congruence. The numbers \( n-b+1, \ldots, n \) form a complete residue system \( \mod b \), there therefore exists a number \( y \) with \( n-b < y \leq n \), satisfying the congruence (6.2). The congruence can be written as an equation

\[ ay + 1 = bx, \]

this gives the desired \( x \). The pair \( x, y \) is easily seen to have the following properties: \( y > 0 \), \( (x, y) = 1 \), \( 0 < x \leq y \). The last relation follows from (6.1) if it is noted that \( a < b \), this is true, as \( \frac{a}{b} \) is not the last element of the Farey Series. Using (6.1) we then have

(6.3) \[ 1 > \frac{x}{y} = \frac{a}{b} + \frac{1}{by} > \frac{a}{b}. \]

It is now sufficient to show that \( \frac{x}{y} = \frac{a'}{b'} \), for then \( x = a', y = b', ba' - ab' = 1 \) and \( b + b' > b + n - b = n \).

Assume \( \frac{x}{y} \neq \frac{a'}{b'} \). Now \( \frac{x}{y} \) is a member of the Farey Series belonging to \( n \), as \( (x, y) = 1 \), \( 0 < x \leq y \leq n \). Further \( \frac{a'}{b'} \) is the right neighbor of \( \frac{a}{b} < \frac{x}{y} \), hence \( \frac{a'}{b'} < \frac{x}{y} \).

This gives

\[ \frac{x}{y} - \frac{a'}{b'} = \frac{xb' - a'y}{yb'} > 0, \]

and therefore
\[ xb' - a'y > 0, \]
the left hand side is an integer, hence
\[ xb' - a'y \geq 1. \]

We then have
\[ \frac{x}{y} - \frac{a'}{b'} \geq \frac{1}{yb'}. \]

For similar reasons
\[ \frac{1}{b'} - \frac{a}{b} \geq \frac{1}{bb'}. \]

Adding these two expressions we get
\[ \frac{x}{y} - \frac{a}{b} \geq \frac{1}{yb'} + \frac{1}{bb'} = \frac{b+vb'}{yb'} > \frac{n}{ybb'} \geq \frac{1}{by}. \]

This contradicts (6.3), hence
\[ \frac{x}{y} = \frac{a'}{b'}, \]

and the theorem is proved.

**Definition 6.2.** Let \( \frac{a}{b}, \frac{a'}{b'} \) be two consecutive members of the Farey Series belonging to \( n \), the expression
\[ \frac{a + a'}{b + b'} \]
is then called the mediant of the two members.

**Theorem 6.2.** The mediant \( \frac{a + a'}{b + b'} \) lies between \( \frac{a}{b} \) and \( \frac{a'}{b'} \), its distance from \( \frac{a}{b} \) and \( \frac{a'}{b'} \) is \( \frac{1}{b(b+b')} \) and
\[ \frac{1}{bb'(b+b')}, \]
respectively.

**Proof.** Let \( \frac{a}{b} < \frac{a'}{b'} \), then it follows from the previous theorem that
\[
\frac{a'}{b'} - \frac{a+a'}{b+b'} = \frac{b a' - ab'}{b' (b+b')} = \frac{1}{b' (b+b')} > 0,
\]
and similarly
\[
\frac{a+a'}{b+b'} - \frac{a}{b} = \frac{1}{b' (b+b')} > 0.
\]
It follows from this theorem, that the mediants are not members of the Farey Series.

**Theorem 6.3.** Every real number \(0 \leq \xi \leq 1\) can be written as
\[
\xi = \frac{a}{b} + \gamma,
\]
where
\[(a,b)=1, \; 0 < b \leq n, \; 0 \leq a \leq b, |\gamma| \leq \frac{1}{b(n+1)}.
\]

**Proof.** The line segment \(0 \leq \xi \leq 1\) is divided by the mediants of the Farey Series belonging to \(n\) into certain abutting intervals \(I(\frac{a}{b})\), the intervals can be characterized by that member of the Farey Series that lies in it. Let \(\xi \in I(\frac{a}{b})\), then it follows from the preceding theorem that
\[
|\gamma| = \left| \frac{a}{b} - \xi \right| \leq \frac{1}{b(n+1)}, \frac{1}{b(n+1)}.
\]
where \(a', b' < \frac{a}{b} < a_b\) are consecutive members of the Farey Series belonging to \(n\). It follows from Theorem 6.1 that
\[
b+b' \geq n+1, \; b+b' \geq n+1.
\]
Hence \(|\gamma| \leq \frac{1}{b(n+1)}\).

**Section 7.**

**Lemma A (Hardy and Littlewood)**

From this section on we shall denote by \(k\) a fixed integer \(k \geq 3\), we further agree to write \(s=4k\). We now introduce the following functions of \(k\) and \(p\).
\( t = t(k,p) \) is to be the number, that satisfies the two conditions \( p^t \not\equiv k \mod p^{t+1} \).

The other expressions are defined as follows:

\[ b = \begin{cases} 
  t+1 & \text{for } p=2, \\
  t+\lambda & \text{for } p>2,
\end{cases} \]

\[ w = p^b \]

\[ k_0 = \frac{k}{p^t}, \text{ hence } p \not\mid k_0. \]

**Definition 7.1.** Let \( \phi \) be a primitive unit root of order \( q \) (i.e. \( \phi = e^{2\pi i q} \), where \((a,q) = 1\)). We then define

\[ S_\phi = S(k) = \sum_{m=1}^{q-1} \phi^m. \]

As \( m_1 \equiv m_2 \pmod{q} \), if \( m_1 \equiv m_2 \pmod{q} \), and as \( \phi^m \equiv \phi^n \pmod{q} \), we have

\[ S_\phi = \sum_{m} \phi^m, \]

where \( m \) runs over an arbitrary complete residue system \( \pmod{q} \).

**Definition 7.2.** We now define

\[ a_q = A_q(k,n) = \sum_{\phi} (S_\phi)^s \phi^{-n}, \]

where \( \phi \) runs over all \( \phi(q) \) primitive unit roots of order \( q \).

**Theorem 7.1.** For \( q > 1 \)

\[ \sum_{n=0}^{q-1} e^{2\pi i n \frac{n}{q}} = 0. \]

**Proof.** Birkhoff's principle of sufficient reason could be used. It also can be proved as follows.
\[
\sum_{n=0}^{q-1} e^{2\pi i n q} = \sum_{n=0}^{q-1} e^{2\pi i n+1 q} = e^{2\pi i \frac{1}{q}} \sum_{n=0}^{q-1} e^{2\pi i n}. 
\]

Hence
\[
(1 - e^{2\pi i \frac{1}{q}}) \sum_{n=0}^{q-1} e^{2\pi i n} = 0. 
\]

Therefore, we must have \(\sum e^2 i n = 0.\)

**Theorem 7.2.** \(A_1 = 1.\)

**Proof.** \(S_1 = 1.\) Hence \(A_1 = (\frac{S}{1})^{1-n} = 1.\)

**Theorem 7.3.** \(A_q\) is real.

**Proof.** If \(\varphi\) runs over all primitive \(q\) th unit roots, so does \(\varphi^{-1}.\) Further \(\varphi = \overline{\varphi}, \overline{S} = S\overline{\varphi} = S\varphi\) and \(\overline{q} = q,\) this gives
\[
A_q = \sum_{\varphi} (\overline{\varphi^q})^S (\overline{\varphi}^{-1})^{-n} = \sum_{\varphi} (\overline{\varphi^q})^S \overline{\varphi}^{-n} = \overline{A_q}. 
\]

Now \(A_q = \overline{A_q}\) if and only if \(A_q\) is real.

**Theorem 7.4.** If \(k = 4, p > 2\) or if \(k \neq 4, p \geq 2\) then \(k \geq 2t + 1.\)

**Proof.** We consider four cases.

1) \(t = 0,\) then \(k \geq 3 > 1 = 2t + 1.\)
2) \(t > 1, p > 2,\) then \(k \geq p \geq 3 \geq 2t + 1.\)
3) \(1 \leq t \leq 2, p = 2,\) then \(k \geq 6 > 2t + 1.\) We have \(k \geq 6,\) because of \(k = 2m, k \geq 3, k \neq 4.\)
4) For \(t > 2, p = 2,\) we have \(k \geq 2t > 2t + 1.\)

Cases 1) and 2) take care of \(p > 2,\) and cases 1), 3) and 4) of \(p = 2, k \neq 4.\)

**Theorem 7.5.** \(k \geq b.\)

**Proof.** We again consider four cases.

1) \(p > 2,\) then it follows from Theorem 7.4 that \(k \geq 2t + 1 \geq t + 1 = b.\)
2) \( p=2, \ k\neq 4, \ t>0, \) then Theorem 7.4 gives \( k \geq 2t+1 \geq t+2 = b. \)

3) \( p=2, \ k\neq 4, \ t=0, \) then \( k \geq 3 > 2 = b. \)

4) \( p=2, \ k=4, \) then \( t=2, \) that is \( k=t+2=b. \)

**Theorem 7.6.** For \( h>1, \)

\[
(x+y p^{h-1})^k \equiv x^k + k x^{k-1} y p^{h-1} \pmod{p^h}.
\]

**Proof.** The binomial expansion of the left hand side gives the first two terms on the right. All other terms of the expansion are congruent 0 mod \( p^h, \) as they contain the factor \( p^{2(h-1)}. \) (\( 2h-2 \geq h \) for \( h>1).\)

**Theorem 7.7.** For \( h>k, \)

\[
(x+y p^{h-1-t})^k \equiv x^k + k_0 x^{k-1} y p^{h-1} \pmod{p^h}.
\]

**Proof.** We first note that \( h-1-t \geq k-t \geq b-t > 0. \)

The first two terms of the binomial expansion of the left hand side are to be found on the right hand side, because we have

\[
k x^{k-1} y p^{h-1-t} = k_0 x^{k-1} y p^{h-1} \quad (k_0 = \frac{k_t}{p}).
\]

It remains to show that for \( 2 \leq r \leq k \)

\[
\binom{k}{r} p^{r(h-1-t)} \equiv 0 \pmod{p^h}.
\]

For \( k=4, p=2 \) we have \( t=2, \ h \geq 5 \) and it is easily seen that the terms \( 6, 2^2(h-3), 4, 2^3(h-3), 2^4(h-3) \) are divisible by \( p \cdot 2^h. \) For the remaining cases we can use Theorem 7.4, which leads to

\[
r(h-1-t) \geq 2(h-1-t) \geq 2(h-1-k) - \frac{k-1}{2} = h + (h-1-k) \geq h.
\]
Theorem 7.8. If either \( t=0, \ h>1 \) or \( t>0, \ h>k \), and if \( \rho \) is a primitive \( p^h \) th unit root, then
\[
\sum_{z=0}^{p^{h-1}} \rho^z = 0
\]

Proof. In both cases we have \( h>t+1 \), we therefore can write
\[
\sum_{z=0}^{p^{h-1}} \rho^z = \sum_{x=0}^{p-1} \rho^x \sum_{y=0}^{p^{h-1}-1} \rho^{(x+y)p^{h-1}-1} k
\]
as any \( z, \ p \neq z, \ 0 \leq z \leq p^{h-1} \) can be expressed uniquely as
\[
z = x + yp^{h-1-t},
\]
where \( 0 \leq x \leq p^{h-1-t}-1, \ p \neq x, \) and \( 0 \leq y \leq p^{t-1}-1. \)

Now we apply Theorem 7.6 in case \( t=0, \ h>1 \) (note, for \( t=0 \) \( k=k_0 \)) and Theorem 7.7 in the other case, hence we get
\[
\sum_{z=0}^{p^{h-1}} \rho^z = \sum_{x=0}^{p-1} \rho^x \sum_{y=0}^{p^{h-1}-1} \left( \rho^{k_0x+k-1p^{h-1}} \right)^y
\]
Now \( p \neq k_0, \ p \neq x \) hence \( p \neq k_0x^{k-1} \) and \( \rho^{k_0x+k-1p^{h-1}} \)
is a primitive \( p^h \) th unit root. As \( y \) runs from 0 to \( p^{t+1}-1 \) it runs over \( p^t \) complete residue systems mod \( p \). Hence we have
\[
\sum_{y=0}^{p^{h-1}-1} \left( \rho^{k_0x+k-1p^{h-1}} \right)^y = p^{t} \sum_{y=0}^{p^{h-1}-1} \left( \rho^{k_0x+k-1p^{h-1}} \right)^y
\]
According to Theorem 7.1 the sum on the right hand side is equal to zero for all \( x, p \neq x \), and hence the theorem follows.

**Theorem 7.9.** Let \( h > k \) and let \( \varphi \) be a primitive \( p^h \) th unit root, then

\[
S_\varphi = p^{k-1}p_k.
\]

**Proof.** We have

\[
S_\varphi = \sum_{z=0}^{p^{k-1}} \varphi^z - \sum_{z=0}^{p^{k-1}} \varphi^z.
\]

According to Theorem 7.8 the second sum of the last expression is equal to zero, hence we have

\[
S_\varphi = \sum_{z=0}^{p^{k-1}} \varphi^z = \sum_{n=0}^{p^{k-1}} \varphi^{(p-1)n}.
\]

Now \( \varphi^{p^k} \) is a primitive \( p^{h-k} \) th unit root, and as \( u \) runs from 0 to \( p^{h-1}-1 \), it covers \( p^{k-1} \) complete residue systems mod \( p^{h-k} \), hence

\[
S_\varphi = p^{k-1} \sum_{n=0}^{p^{k-1}} (\varphi^{p^k})^n = p^{k-1}S_\varphi.
\]

**Theorem 7.10.** Let \( p \neq k \), \( 2 \leq h \leq k \), and let \( \varphi \) be a primitive \( p^h \) th unit root, then

\[
S_\varphi = p^{h-1}.
\]

**Proof.** The first case of Theorem 7.8 \((t=0, h > 1)\) applies here, hence we have

\[
S_\varphi = \sum_{z=0}^{p^{k-1}} \varphi^z + 0 = \sum_{z=0}^{p^{k-1}} \varphi^z = p^{k-1}.
\]
Here \( p^k z = 1 \) as \( p/z \), \( k > h \) and hence \( p^h / z^k \). It is then sufficient to remark, that there are exactly \( p^{h-1} \) numbers \( z \) satisfying \( 0 \leq z \leq p^{h-1}, p/z \).

**Theorem 7.11.** Let \( \chi(a) \) be an character mod \( p \) but not a main character (we do not know as yet whether such a character exists, but shall construct one later.)

Let \( \rho \) be a primitive \( p \) th unit root, then

\[
\left| \sum_{\omega=1}^{p-1} \chi(\omega) \rho^a \right| = \sqrt{p}
\]

**Proof.** We have \( |a|^2 = a \bar{a} \), hence

\[
\left| \sum_{\omega=1}^{p-1} \chi(\omega) \rho^a \right|^2 = \sum_{\omega=1}^{p-1} \chi(\omega) \rho^a \sum_{f=1}^{p-1} \bar{\chi}(f) \rho^{-f}
\]

For \( 1 \leq a \leq p-1 \), \( (a,p)=1 \), then as \( d \) goes through a reduced residue system mod \( p \), so does \( ad \). Now \( \rho^a = \rho^b \) if \( a \equiv b \) (mod \( p \)), further \( \bar{\chi}(a) \) is a character if \( \chi(a) \) is one, and for characters we have \( \chi(a) = \chi(b) \) if \( a \equiv b \) (mod \( p \)). From all this it follows that the second sum in our expression remains unchanged, if the summation is over any arbitrary reduced residue system modulo \( p \). We can therefore write:

\[
\left| \sum_{\omega=1}^{p-1} \chi(\omega) \rho^a \right|^2 = \sum_{\omega=1}^{p-1} \chi(\omega) \rho^a \sum_{d=1}^{p-1} \bar{\chi}(ad) \rho^{-ad}
\]

\[
= \sum_{d=1}^{p-1} \bar{\chi}(ad) \sum_{d=1}^{p-1} |\chi(\omega)|^2 \rho^{a(1-d)}
\]
We here made use of the fact that $\chi(a \cdot d) = \chi(a) \cdot \chi(d)$.

We now recall, that $|\chi(a)| = 1$, hence our expression becomes

$$\sum_{d=1}^{p-1} \chi(d) \sum_{a=1}^{p-1} g^{\chi(a-1-\chi(d))}$$

Now as $\chi(a)$ is not a main character, $\chi(a)$ is not, so using Theorem 5.5 we have

$$\sum_{d=1}^{p-1} \chi(d) = \sum_{d=0}^{p-1} \chi(d) - \chi(0) = 0 - 0$$

We thus can write

$$\sum_{d=1}^{p-1} \chi(d) \sum_{a=1}^{p-1} g^{\chi(a-1-\chi(d))} = \sum_{d=1}^{p-1} \chi(d) \sum_{a=0}^{p-1} g^{\chi(a-1-\chi(d))} - \sum_{d=1}^{p-1} \chi(d) g^0$$

Then Theorem 7.1 leads to

$$\sum_{a=0}^{p-1} g^{\chi(a-1-\chi(d))} = \begin{cases} 0 & \text{for } d \neq 1 \\ p & \text{for } d = 1 \end{cases}$$

Now $\chi(1) = 1$, hence we finally obtain

$$\left| \sum_{a=1}^{p-1} \chi(a) g^{\chi(a)} \right|^2 = \sum_{d=1}^{p-1} \chi(d) \sum_{a=1}^{p-1} g^{\chi(a-1-\chi(d))} = p \chi(1) = p$$

**Theorem 7.12.** If $\chi(a)$ is the main character mod $p$ and if $\rho$ is a primitive $p$th unit root, then

$$\sum_{a=1}^{p-1} \chi(a) \rho^a = -1$$

**Proof.** $\chi(a) = 1$ for $1 \leq a \leq p-1$. Hence

$$\sum_{a=1}^{p-1} \chi(a) \rho^a = \sum_{a=1}^{p-1} \rho^a - g^0 = 0 - 1.$$
p \neq a, so that there exists an x such that
\[ a \equiv g^x \pmod{p^h}, \quad 0 < x \leq \varphi(p^h). \]

Then
\[ a \equiv v^k \pmod{p^h} \]
if and only if \((k, \varphi(p^h)) \mid x.\)

**Proof.** The existence of the numbers \(g\) and \(x\) is insured by Theorem 2.27 and Theorem 2.28 respectively. We first show that
\[ a \equiv v^k \pmod{p^h} \]
implies \(p \nmid v\), that is \(v \equiv g^y \pmod{p^h}\).
Assume the contrary, that is \(v = pr\), that implies
\[ a - p^kr^k = bp^h \]
from which it follows that \(p \nmid a\), this however contradicts our assumption. It remains to solve the congruence
\[ a \equiv g^x = (g^y)^k \equiv (\pmod{p^h}) \]
for \(y.\) Theorem 2.23 states that this congruence is solvable if and only if \(y\) satisfies the congruence
\[ x \equiv ky \pmod{\varphi(p^h)}\].
According to Theorem 2.20 this congruence has solutions if and only if \((k, \varphi(p^h)) \mid x.\)

**Theorem 7.14.** Let \(p > 2, 0 < h \leq t+1,\)
then the number of \(k\)th power residue classes \(\pmod{p^h}\) (prime to \(p^h\)) is equal to \(\frac{p-1}{(k, p-1)}\).

**Proof.** According to the preceding theorem
The number \( a \) belongs to a \( k \) th power residue class, if and only if the number \( x \) is a multiple of 
\( (\varphi(p^k), k) \) and satisfies the inequality \( 0 < x \leq (p^h) \), where, as before, \( x \) is defined by the congruence 
\( a \equiv g^x \pmod{p^h} \). The number of possible \( x \), and hence \( k \) th power residue classes, is therefore equal to 
\[ \frac{\varphi(p^h)}{(k, \varphi(p^h))} = \frac{p^h - 1}{p^h - 1} \frac{p^h - 1}{p^h - 1} = \left( \frac{p - 1}{k_0, p - 1} \right) \]
The last equality follows from the fact that \( p - 1 \neq p \), so that 
\[ (p^{t-h-1}, k_0, p - 1) = (k_0, p - 1). \]

**Theorem 7.15.** Let \( p > 2 \) and let \( g \) be a primitive root mod \( p \). The following \((k, p - 1)\) functions 
\( \chi_m(a), 0 \leq m < (k, p - 1) \) are different characters mod \( p \):

\[ \chi_m(a) = \begin{cases} 
0 & \text{for } p \not| \ a, \\
\frac{\chi_{m^x}}{\chi(x^m)} & \text{for } a \equiv g^x \pmod{p}, 0 \leq x < p - 1.
\end{cases} \]

**Proof.** The four conditions of Definition 5.5, 1 are fulfilled. This is evident for i), ii), iv). For condition iii) we note that 
\[ a_1 \equiv g^{x_1}, \ a_2 \equiv g^{x_2} \pmod{p} \]
implies 
\[ a_1 a_2 \equiv g^{x_1 + x_2} \pmod{p}, \]
and if 
\[ a_1 a_2 \equiv g^{x_3} \pmod{p} \]
we must have 
\[ x_1 + x_2 \equiv x_3 \pmod{p - 1}, \]
and hence by Theorem 2.9
\[ x_1 + x_2 \equiv x_3 \pmod{(k,p-1)} \]
and finally
\[ \chi_m(a_1)\chi_m(a_2) = \chi_m(a_1a_2). \]
Hence the \( \chi_m \) are characters, that they are different follows from the fact that for \( m \neq n \) \( \chi_m(g) \neq \chi_n(g) \).
Evidently \( \chi_0 \) is the main character.

**Theorem 7.16.** Let \( p > 2, a > 0 \), then
\[
\sum_{m=0}^{(k,p-1)-1}\chi_m(a) = \begin{cases} (k,p-1) & \text{for } p \neq a, a \equiv v^k \pmod{p} \\ 0 & \text{otherwise.} \end{cases}
\]

**Proof.** For the first case we evidently have \( \chi_m(a) = 1 \) for all \( m \). This follows from Theorem 7.13, from the same theorem we deduce that \( \chi_m(k,p-1) \neq x \), where \( a \equiv g^x \pmod{p} \). Now let \((k,p-1),x) = d < (k,p-1)\), then \( \sum_m \chi_m(a) \) is equal to \( d \) times the sum of all \( (k,p-1)/d \) unit roots of that order, and hence equal to zero.

**Theorem 7.17.** Let \( p \geq 2 \). If \( h \) runs from 1 to \( p-1 \), then \( h^k \) runs mod \( p \) through all \( k \) th power residue classes \((\neq 0 \pmod{p})\) mod \( p \), there are \( (k,p-1) \) numbers \( h^k \) in every class.

**Proof.** That all classes \( a \) are represented is clear. From Theorem 2.31 it follows that there are \( (k,p-1) \) solutions of the congruence \( x^k \equiv 1 \pmod{p} \).
There are therefore at least \((k, p-1)\) solutions of the congruence

\[ x^k \equiv v^k \equiv a \pmod{p}. \]

Now there are \(\frac{p-1}{(k, p-1)}\) different classes \(a\), this can be seen from Theorem 7.14, with \(h = 1\), if it is noted that \((k, p-1) = (k_0, p-1)\). There are \(p-1\) numbers \(h^k\), hence there are exactly \((k, p-1)\) numbers \(h^k\) in every residue class of \(k\) th power \((\neq 0 \pmod{p})\).

**Theorem 7.18. Let** \(p > 2\), and let \(\rho\) be a primitive \(p\) th unit root. Then

\[ |S_\rho| < k \sqrt{p}. \]

**Proof.** We have

\[ S_\rho = 1 + \sum_{a=1}^{p-1} \frac{a}{\mathbb{Z}} \rho^a = 1 + \sum_{\alpha=1}^{(k, p-1)-1} \frac{\alpha^{(k, p-1)-1}}{\alpha=1} \sum_{m=0}^{\alpha \chi_m(a)} \]

\[ = 1 + \sum_{m=0}^{(k, p-1)-1} \frac{\sum_{\alpha=1}^{(k, p-1)-1} \sum_{\alpha=1}^{\alpha \chi_m(a)}} {\sum_{\alpha=1}^{(k, p-1)-1} \sum_{\alpha=1}^{\alpha \chi_m(a)}} \rho^a \]

This last step is justified by

\[ \sum_{\alpha=1}^{(k, p-1)-1} \chi_0(a) \rho^a = \sum_{\alpha=1}^{(k, p-1)-1} \chi_1(a) \rho^a = -1 \quad \text{(Theorem 7.12)} \]

We now use Theorem 7.11 and obtain

\[ |S_\rho| \leq \sum_{m=0}^{(k, p-1)-1} \left| \sum_{\alpha=1}^{(k, p-1)-1} \chi_m(a) \rho^a \right| \leq (k, p-1) \sqrt{p} < k \sqrt{p}. \]
Theorem 7.19. Let \((q_1, q_2) = 1\), let \(\rho_1\) and \(\rho_2\) be primitive unit roots of degree \(q_1\) and \(q_2\) respectively, then

\[ S_{\rho_1 \rho_2} = S_{\rho_1} S_{\rho_2}. \]

Proof. The root \(\rho_1 \rho_2\) is a primitive \(q_1 q_2\) th unit root. We have

\[ \rho_1 = e^{2\pi i q_1^{-1}}, \quad \rho_2 = e^{2\pi i q_2^{-1}}, \quad (\alpha_1 \alpha_2) = (a_2, q_2) = 1 \]

\[ S_{\rho_1 \rho_2} = e^{2\pi i \frac{a_1 q_2 + a_2 q_1}{q_1 q_2}} \]

and \((a_1 q_2 + a_2 q_1, q_1 q_2) = 1\). Now by Theorem 2.15 the number \(z q_1 y q_2\) runs over a complete residue system mod \(q_1 q_2\) if \(z\) and \(y\) run over complete residue systems mod \(q_1\) and mod \(q_2\) respectively. Hence

\[ S_{\rho_1 \rho_2} = \frac{q_1}{q_2} \sum_{y=1}^{q_2} (\rho_1 \rho_2)^{(z q_1 + y q_2) k} \]

Now

\[ (z q_1 + y q_2)^k \equiv (z q_1)^k + (y q_2)^k \pmod{q_1 q_2} \]

hence

\[ S_{\rho_1 \rho_2} = \sum_{y=1}^{q_2} (\rho_1 \rho_2)^{(z q_1)^k} + (y q_2)^k \]

\[ = \sum_{y=1}^{q_2} \rho_1^{(y q_2)^k} \sum_{z=1}^{q_1} \rho_2^{(z q_1)^k} = \sum_{y=1}^{q_2} \rho_1^{y k} = S_{\rho_1} S_{\rho_2} \]

We made here use successively of the facts that

\[ \rho_1 q_1 = 1, \rho_2 q_2 = 1, \]

and that if \(y\) and \(z\) run through complete residue systems mod \(q_1\) and mod \(q_2\) respectively, the same is true for \(y q_2\) and \(z q_1\) (because \((q_1, q_2) = 1\)).
Theorem 7.20. Let \((q_1, q_2) = 1\), then (\(n\) being fixed)

\[ \mathcal{A}_{q_1 q_2} = \mathcal{A}_{q_1} \mathcal{A}_{q_2} \]

Note: This theorem will not be needed until the next section.

**Proof.** From Theorem 2.16 it follows that, if \(\rho_1\) runs through all primitive \(q_1\) th unit roots, and \(\rho_2\) runs through all primitive \(q_2\) th unit roots, then \(\rho_1 \rho_2\) runs through all primitive \(q_1 q_2\) th unit roots. We further make use of the preceding theorem, this gives

\[ R_{q_1 R_{q_2}} = (q_1, q_2)^{-1} \sum_{\rho_1, \rho_2} (S_{\rho_1}, S_{\rho_2})^5 (\rho_1, \rho_2)^{-n} = R_{q_1, q_2} \]

**Definition 7.3.** Let \(\rho\) be a primitive \(q\) th unit root, then we set

\[ \tau = q^{\frac{1}{h}} - 1 \]

\[ \tau \rho = q^{\frac{1}{h}} S_{\rho} \]

If in particular \(q = p^h\), we have

\[ \tau \rho = p^{\frac{h}{h}} S_{\rho} \]

**Theorem 7.21.** If \((q_1, q_2) = 1\) and \(\rho_1\) and \(\rho_2\) are primitive \(q_1\) th and \(q_2\) th unit roots respectively, then

\[ \tau \rho_1 \rho_2 = \tau \rho_1 \tau \rho_2 \]

**Proof.** Using Theorem 7.19, we have

\[ \tau \rho_1 \rho_2 = (q_1, q_2)^{\frac{1}{h}} S_{\rho_1 \rho_2} = (q_1^{\frac{1}{h}} S_{\rho_1})(q_2^{\frac{1}{h}} S_{\rho_2}) \]

\[ = \tau \rho_1 \tau \rho_2 \]
Theorem 7.22. Let \( \rho \) be a primitive \( q \) th unit root. Let \( p \equiv q \) and let \( h(q,p) \geq 1 \) be such that \( p^h \equiv q \). Then with suitable choice of the primitive \( p^h(q,p) \) th unit roots \( \rho(p^h(q,p)) \), we have
\[
T_{\rho} = \prod_{p \equiv q} T_{\rho(p^h(q,p))}.
\]

Proof. Successive applications of the preceding theorem yield this result. The \( \rho(p^h(q,p)) \) can be so chosen as to yield the desired
\[
\rho = \prod_{p \equiv q} \rho(p^h(q,p)).
\]
For let \( \rho(p^h(q,p)) \) be arbitrary roots, then
\[
\prod \rho(p^h(q,p)) = e^{\frac{2\pi i}{q}} (\frac{a}{q}q) = 1
\]
where \( (\frac{a}{q}) = 1 \). Let \( \phi = 1 \) be the desired value, then by Theorem 2.19 there exists a \( d \) such that \( (d,q) = 1 \) and
\[
ad \equiv b \pmod q.
\]
The \( d \) th powers of the originally chosen \( \rho(p^h(q,p)) \) then combine to give the desired \( \rho \).

Theorem 7.23. If \( q=p^h \), \( h \equiv k \) and if \( \rho \) is a primitive \( q \) th unit root, then
\[
T_{\rho} = T_{\rho^k}.
\]

Proof. Using Theorem 7.9 we have
\[
T_{\rho} = p_{\frac{k}{h}} - k \cdot S_{\rho} = p_{\frac{k}{h}} - k (p_{k-1} S_{\rho^k})
\]
\[
= p_{\frac{k-h}{h}} \cdot S_{\rho^k} = T_{\rho^k} \quad (S_{\rho^k} \text{ of degree } k-h)
\]

Theorem 7.24. For \( q=p^h \), \( h>0 \) and every unit root of degree \( q \), we have
\[
|T_{\rho}| \leq \left\{ \begin{array}{ll}
1 & \text{for } p \geq k^5 \\
k & \text{always}
\end{array} \right.
\]
Proof. Theorem 7.23 allows us to assume, without loss of generality, that \(1 \leq h \leq k\). That is to any given \(T_p\) we apply the theorem \(n\) times, where \(n\) is such that \(k \geq h - nk \geq 1\). We now consider three cases:

1) If \(p \not| k\), we need only the following rough approximation

\[ |S_p| = \left| \frac{p^h}{x=1} x^k \right| \leq p^h \]

for then

\[ |T_p| = p^{\frac{h}{k}} - 1 |S_p| \leq p^{\frac{h}{k}} - 1 \leq 1 \]

2) If \(p \not| k\), \(2 \leq h \leq k\), Theorem 7.10 can be used, this gives

\[ S_p = p^{h-1} \]

and hence

\[ T_p = p^{\frac{h}{k}} - 1 |S_p| = p^{\frac{h}{k}} - 1 p^{h-1} = p^{\frac{h}{k}} - 1 \leq 1 \]

3) If \(p \not| k\), \(h=1\), then Theorem 7.18 gives \((p>2)\)

\[ |S_p| < \frac{k}{\sqrt[p]{p}} \]

The same relation evidently holds for \(p=2\), thus we have

\[ |T_p| = p^{\frac{1}{k}} - 1 |S_p| < p^{\frac{1}{k}} - 1 p^{\frac{1}{k}} \leq \frac{k}{\sqrt[p]{p}} \leq \left\{ \begin{array}{ll} 1, & p \geq k^6 \\ h, & \text{always} \end{array} \right. \]

If \(p \geq k^6\), either case 2) or 3) applies.

Theorem 7.25.

\[ |T_p| \leq c_4(k) \]

Proof. From Theorem 7.22 we have

\[ |T_p| = \prod_{p|q} |T_p'(p^h(a, p))| \]
to this we apply Theorem 7.24, thus we obtain
\[ |T_\ell| \leq \prod_{p|\ell} \frac{1}{p} \prod_{p \leq \ell} 1 \leq \ell^{k^6} = C_4(k) \]

Theorem 7.25 (Lemma A.)
\[ |S_\ell| < C_4(k) q^{-\frac{1}{k}} \]

Proof. This is an immediate consequence of Theorem 7.25 and Definition 7.3.

Section 8.

Lemma B (Hardy and Littlewood)

Theorem 8.1. \( \exists q > 0 \).
\[ |A_q| \leq C_{20}(k) q^{-2} \]

Proof. According to Theorem 7.26 we have
\[ |A_q| \leq \sum_{q} \left| \frac{S_{\ell}}{q} \right|^{4,k} \leq C_1 \left( C_4(k) q^{-\frac{1}{k}} \right)^{4,k} = C_4(k) q^{-3} \]

Theorem 8.2.
\[ \sum q^{-3} \text{ converges.} \]

Proof. \( \sum q^{-3} \text{ converges, hence Theorem 8.1 gives the desired result.} \)

Definition 8.1.
\[ \mathcal{T} = \mathcal{T}(k, \eta) = \sum_{q=1}^{\infty} A_q(k, \eta). \]

From now until the end of the chapter we shall denote by \( p_h \) the \( h \) th prime number.
Theorem 8.3.

\[ \mathcal{G} = \lim_{m \to \infty} \sum_{q | m} A_q \]

Proof. Let us set

\[ \sum_m = \sum_{q | m} A_q \]

then evidently

\[ | \mathcal{G} - \sum_m | \leq \sum_{q=m}^\infty | A_q | = R_m \]

Now from Theorem 8.2 it follows that \( R_m \to 0 \) as \( m \to \infty \)
and therefore

\[ \mathcal{G} = \lim_{m \to \infty} \sum_m \]

Theorem 8.4.

\[ \sum_{q | p} A_q > 1 - \frac{1}{p^2}, \text{ if } m > 0, \ p > c_{21}(k). \]

Proof. We first recall that according to Theorem 7.3

\( A_q \) is real. We also showed (Theorem 7.2) that \( A_1 = 1 \), thus

we have

\[ \sum_{q | p} A_q \geq R_1 - \sum_{f=1}^\infty | A_{pf} | > 1 - c_{20}(k) \sum_{f=1}^\infty p^{-3f} \]

The last step is justified by Theorem 8.1, then

\[ \sum_{q | p} A_q \geq 1 - \frac{c_{20}(k)}{p^3 - 1} > 1 - \frac{2 c_{20}(k)}{p^3} > 1 - \frac{1}{p^2} \]

for \( p > 2c_{20}(k) = c_{21}(k) \).

Theorem 8.5. The congruence

\[ \sum_{m=1}^{4k-1} x_m^k \equiv n \pmod{w} \]

is solvable for every \( n \), \( 1 \leq n \leq w-1 \).

Note: This is a "Waring Theorem" for congruences of a special type.
Proof. We consider two cases:

1) \( p = 2 \), then \( 1 \leq n \leq w-1 = 2^b - 1 = 2^{t+2} - 1 = 4 \cdot 2^t - 1 \leq 4k - 1 \) (\( 2^t / k \) by definition). Hence we have

\[
n = \sum_{m=1}^{N} 1^k + \sum_{m=n+1}^{w} 0^k.
\]

2) \( p > 2 \), then \( w = p^b = p^{t+1} \). We divide the numbers \( n \) \( 1 \leq n \leq w-1 \) into classes. \( n_1 \) and \( n_2 \) belong to the same class, if there exists an \( x \) such that

\[
x^k n_1 \equiv n_2 \pmod{w}, \quad p \neq x.
\]

The equivalence relationship \( (n_1 \sim n_2) \) thus defined is evidently reflexive and transitive, it also is symmetric. This we see as follows, let \( k = p^t b, \ \varphi(w) = p^t (p-1) \), then using Fermat's Theorem we have

\[
n_1 \equiv x^b \varphi(w) n_1 \equiv x^b \varphi(w) - k n_2 \equiv x^b p^t (p-1) n_2
\]

\[
\equiv (x^{p-2})^k n_2 \pmod{w}, \quad p \neq x^{p-2}.
\]

We next note that \( (n, w) \) has the same value \( p^d \) for all \( n \) in one class, this can be seen as follows:

\( p^{t+1} / (x^k n_1 - n_2) \) and \( p \neq x \), hence if \( p^d = (w, n_2) \), then

\( p^d / n_1, \ p^d / (w, n_1) \). Now assume \( p^{d+1} / (w, n_1) \), then the same argument turned around would give \( p^{d+1} / (w, n_2) \) which is a contradiction, hence \( (w, n_1) = (w, n_2) \) for all \( n_1 \sim n_2 \).

Next we try to find the number of different classes. We first find the number of classes having a fixed \( (w, n) = p^d \).
In every one of these classes let \( n_1 \) be a fixed element, we obtain all elements of one class, if we consider the different numbers \( \equiv x^{k_1} \pmod{p^{t+1}}, p \neq x \).

This is the same (Theorem 2.10) as considering the different numbers \( \equiv x^{k_0} \pmod{p^{t-d+1}}, p \neq x, p \neq b \).

And this finally is equivalent to determining the number of \( k \) th power residue classes, prime to \( p^{t-d+1}, \mod p^{t-d+1} \). This number was found, Theorem 7.14, to be

\[
\frac{p-1}{(k_0, p-1)}.
\]

The number of all elements \( n, 1 \leq n \leq w-1 \), with \( (n, w) = p^d \), is evidently \( \varphi(p^{t-d+1}) = p^{t-d}(p-1) \).

For every \( d \) there therefore are \( p^{t-d}(k_0, p-1) \) different classes. We get the number of all classes by summing over \( d \)

\[
r = (k_0, p-1) \sum_{d=0}^{d} p^{t-d} = \frac{p^{t+1}-1}{p-1} (k_0, p-1)
\]

\[
\frac{p^{t+1}}{p-1} k_0 = \frac{pk_0}{p-1} \leq \frac{3}{2} k < 4k,
\]

hence

\[
r \leq 4k-1.
\]

Now let \( h(n), 1 \leq n \leq w-1 \), be the smallest \( v \) for which the congruence

\[
\sum_{m=1}^{\varphi} x_m^k \equiv n \pmod{w}
\]

is solvable. Evidently \( h(n_1) = h(n_2) \), if \( n_1 \sim n_2 \). \( h(n) \) therefore \( m \) is a function of the class only, and has
therefore at most $4k-1$ values. Now

$$n+1 = n+1^k,$$

and hence

$$h(n+1) \leq h(n)+1.$$

Further $h(1) = 1$, hence

$$h(n) \leq 4k-1 \text{ for } 1 \leq n \leq w-1.$$

Theorem 8.6. The congruence

$$\sum_{m=1}^{4k} x_m^k \equiv n \pmod{w}$$

has one solution, with $x_1 = 1$.

Proof. Without loss of generality we can set

$$1 \leq n \leq w.$$ For $n = 1$, we have

$$n \equiv 1^k + \sum_{m=2}^{4k} a^k \pmod{w}.$$ For $2 \leq n \leq w$, we know from the preceding theorem that

$$n-1 \equiv \sum_{m=2}^{4k} x_m^k \pmod{w}$$

is solvable, hence

$$n \equiv 1^k + \sum_{m=2}^{4k} x_m^k \pmod{w}$$

is solvable.

Definition 8.2. For $z > 0$ we define as

$$\mathcal{M}(z) = \mathcal{M}(z,k,n)$$

the number of solutions of the congruence

$$\sum_{m=1}^{z} x_m^k \equiv n \pmod{z}.$$

Theorem 8.7.

$$\sum_{q|z} q \mathcal{A}_q = z - 1 - \mathcal{M}(z).$$

Proof. If $p$ runs over all $z$ th unit root, then

$$\sum_{q} q^a = \begin{cases} z & \text{for } a \equiv 0 \pmod{z}, \\ 0 & \text{otherwise} \end{cases}$$
In the second case let \((a,z) = d\), then the sum runs \(d\) times over all \(\frac{z}{d}\) th unit roots, and Theorem 7.1 gives the desired result. We therefore can write

\[
(1 - s)^{-s} M(z) = (1 - s)^{-s} \sum_{x_1, \ldots, x_3} \sum_{x_{2m+1}} \sum_{x_m} x_m^z - n
\]

\[
= (1 - s)^{-s} \sum_{x_{2m+1}} x_m^z - n \left( \sum_{x_m} x_m^z \right)^s
\]

Now for every \(q\) let \(q, q \neq z\), be the actual degree of \(\zeta\) (i.e. \(\zeta\) is a primitive \(q\) th unit root), then

\[
\sum_{x_{2m+1}} \zeta^{x_m^z} = \frac{q}{q} \sum_{y=1}^{q} \zeta^{y^z}
\]

because for every \(1 \leq y \leq \frac{q}{z}\) there are exactly \(\frac{z}{q}\) numbers \(x, 1 \leq x \leq z\) with

\[
x \equiv y \pmod{q},
\]

and hence

\[
x^k \equiv y^k \pmod{q}.
\]

We can therefore write

\[
(1 - s)^{-s} M(z) = (1 - s)^{-s} \sum_{\zeta} \zeta^{-n} \left( \frac{q}{q} \sum_{y=1}^{q} \zeta^{y^z} \right)^s
\]

\[
= \sum_{\zeta} \left( \frac{\zeta}{q} \right) \zeta^{-n}
\]

The set of all the \(z\) th unit roots is composed of all primitive \(qz\) th unit roots, for all \(q \neq z\). To prove this we first note that all \(z\) th unit roots are primitive \(q\) th unit roots. Now let \(\zeta_1 = e^{2\pi i \frac{a}{q}}\) be an arbitrary primitive \(q\) th unit root, then

\[
\zeta_1^{2 \pi i \frac{a}{q} (z = qb)}
\]

is contained in the set of all \(z\) th unit roots. It remains to show that \(a\) can
not be both primitive $q_1$ th and primitive $q_2$ th unit root $(q_1 \neq q_2)$. Assume $q_1 < q_2$, $z = b_1q_1 = b_2q_2$, and assume there existed primitive roots of order $q_1$ and $q_2$, both equal to the same $z$ th root, then there would have to exist numbers $a_1$ and $a_2$, with $(a_1, q_1) = 1$, $(a_2, q_2) = 1$, satisfying the congruence

$$a_1 b_1 \equiv a_2 b_2 \pmod{z}.$$  

Multiplying this by $q_1$ and then dividing by $z$, we obtain

$$a_1 = \frac{a_2 b_2 q_1}{b_2 q_2} + r = \frac{a_2 q_1}{q_2} + r,$$

this implies $q_2 \mid q_1$, which is impossible.

The set of all $\mathfrak{D}$ therefore decomposes in the described way, and we can write

$$z^{1 - s}(z) = \sum_{q \mid z} \frac{a_q}{q}.$$  

**Theorem 8.8.** If $z_1 > 0$, $z_2 > 0$ and $(z_1, z_2) = 1$, then

$$\sum_{q \mid z_1 z_2} a_q = \sum_{q \mid z_1} a_q \sum_{q \mid z_2} a_q.$$  

**Proof.** If $(z_1, z_2) = 1$ every $q \mid z_1 z_2$ can be expressed uniquely in the form $q = q_1 q_2$, where $q_1 \mid z_1$, $q_2 \mid z_2$. Hence evidently $(q_1, q_2) = 1$.

Thus we have

$$\sum_{q \mid z_1 z_2} a_q = \sum_{q \mid z_1} \sum_{q \mid z_2} a_{q_1 q_2}$$

An application of Theorem 7.20 then leads to

$$\sum_{q \mid z_1 z_2} a_q = \sum_{q \mid z_1} \sum_{q \mid z_2} a_{q_1} a_{q_2} = \sum_{q \mid z_1} a_{q_1} \sum_{q \mid z_2} a_{q_2}.$$
Theorem 8.9. If $m > 0$, then

$$\sum_{q \mid m} a_q = \prod_{p \leq p_m} \sum_{q \mid p} a_q$$

**Proof.** This is an immediate consequence of the preceding theorem.

Theorem 8.10. If $a \equiv 1 \pmod{m}$, then there exists an $x$ such that

$$x^k \equiv a \pmod{p^m}.$$ 

**Proof.** Let $g = g(p)$ be equal to 5 for $p = 2$, and for $p > 2$ let $g$ be equal to a number which belongs to the exponent $\varphi(p^n) \pmod{p^n}$ for every $n > 0$. (The existence of such a number is insured by Theorem 2.27.) If $n > 1$, $p = 2$, and if $n > 0$, $p > 2$, then belongs to the exponent $\varphi(p^{n-2} + t)$ mod $p^n$. That this is true for $g = 5$, $p = 2$, follows from Theorem 2.29 (recall the definition of $b$).

We now choose $z \geq 0$ in such a way that

$$a \equiv g^z \pmod{p^m}.$$ 

That this is possible for $p > 2$ follows from Theorem 2.28, because we assumed $a \equiv 1 \pmod{w}$, that is certainly $p \not| a$. For $p = 2$ it follows from Theorem 2.30, because we assumed $a \equiv 1 \pmod{w}$, hence a fortiori $a \equiv 1 \pmod{4}$. As $m \geq b$ we have

$$a \equiv g^z \pmod{p^b},$$

and using our assumption

$$g^z \equiv 1 \pmod{p^b}.$$
This implies, (Theorem 2.23)
\[ z \equiv 0 \pmod{p^t(p-1)}. \]

Now
\[ (k(p-1), p^{m-b+t}(p-1)) = p^t(p-1), \]

Hence by Theorem 2.20 we can find a \( \gamma \) such that
\[ z \equiv \gamma k(p-1) \pmod{p^{m-b+t}(p-1)} \]

that is
\[ z \equiv \gamma k(p-1) \pmod{p^{m-1}(p-1)}. \]

Using Theorem 2.23 again, we then have
\[ g^z \equiv g^{\gamma k(p-1)} \pmod{p^m}, \]

and thus
\[ a \equiv (g^{\gamma(p-1)})^k \pmod{p^m}. \]

Theorem 8.11. For \( m \geq b \) we have
\[ \Delta(p^m) \geq p^{(m-b)(s-1)} = p^m(s-1), s = \frac{m}{2}. \]

Proof. According to Theorem 2.16 every residue class \( \pmod{p^b} \) consists of exactly \( p^{m-b} \) residue classes \( \pmod{p^m} \). According to Theorem 8.6, there then exist at least \( p^{(m-b)(s-1)} \) systems \( x_2, \ldots, x_s \pmod{p^m} \) such that
\[ 1 + \sum_{\gamma = 1}^{s} x_\gamma^k \equiv n \pmod{p^b}, 1 \quad x_\gamma \quad p^m. \]

We now choose \( x_1 \) such that
\[ x_1^k \equiv n - \sum_{\gamma = 2}^{s} x_\gamma^k \pmod{p^m}. \]

This is possible, as
\[ n - \sum_{\gamma = 2}^{s} x_\gamma^k \equiv 1 \pmod{w}, \]

by the preceding theorem.
Hence there are at least \( p^{(m-b)(s-1)} \) solutions of the congruence
\[
\sum_{\tau=1}^{s} x_{\tau}^{m} \equiv n \pmod{p^{m}},
\]
and the theorem follows from Definition 8.2.

**Theorem 8.12.** If \( m \geq k \), then
\[
\sum_{q \mid p^{m}} A_q \geq \frac{1}{c_{22}(k,p)}.
\]

**Proof.** Combining Theorem 8.11 with Theorem 8.7 and noting that if \( m \geq k \), then a fortiori \( m \geq b \), we have
\[
\sum_{q \mid p^{m}} A_q = p^{m(1-s)\tilde{\alpha}(p^m)} \geq w^{1-s} = \frac{1}{c_{22}(k,p)}.
\]

**Theorem 8.13.** (Lemma B.)
\[
\mathcal{G}(k,n) > \frac{1}{c_5(k)}.
\]

**Proof.** According to Theorems 8.12, 8.9 and 8.4 we have for all \( m \) with \( p_m \geq c_{21}(k), m \geq k \)
\[
\sum_{q \mid p^{m}} A_q = \prod_{p \leq c_{21}(k)} \sum_{q \mid p^{m}} A_q \bigg/ \prod_{p \leq c_{21}(k)} \left( \frac{1}{c_{22}(k,p)} \right) \prod_{p > c_{21}(k)} \left( 1 - \frac{1}{p^2} \right) \bigg/ \left( \min_{p \leq c_{21}(k), p > c_{21}(k)} \right) \frac{c_{21}(k)}{2} = c_5(k)
\]
This theorem then follows from Theorem 8.3.
Section 9.

**LEMMA C (VINOGRAĐAN).**

A lemma of a similar nature due to Weyl, was used by Hardy and Littlewood, to manage the so called minor arcs. It was largely because of Vinogradow's improvement on Weyl's result, that he was able to obtain so much better values for the bounds of $G(k)$.

**Theorem 9.1.** If $q > 1$, $(a,q) = 1$, $\alpha$ real,

$$\left| x - \frac{a}{q} \right| \leq \frac{1}{q^2}, \gamma \text{ real},$$

then

$$\sum_{x=0}^{q-1} \min(q, \frac{1}{q^2} + \alpha^2) < 8q \log q.$$

($\min(q, \frac{1}{q^2}) = q$).

**Proof.** We consider two cases:

1) $1 < q < 6$, then

$$\sum_{x=0}^{q-1} \min(q, \frac{1}{q^2} + \alpha^2) \leq \sum_{x=0}^{q-1} q = 5q \leq 8q \log q.$$

2) $q \geq 6$. We choose $m$ in such a way that (this is evidently possible)

$$\left| \gamma - \frac{m}{q} \right| \leq \frac{1}{2q},$$

we then have for $0 \leq x \leq q-1$

$$\left| \gamma + \alpha x - \frac{m + \alpha x}{q} \right| = \left| \gamma - \frac{m}{q} + x (\alpha - \frac{\alpha}{q}) \right| < \frac{1}{2q} + q \frac{1}{2q} = \frac{3}{2q}$$

and

$$\gamma + \alpha x = \frac{m + \alpha x + \theta(x)}{q}, \quad \left| \theta(x) \right| < \frac{3}{2q}. $$
Now \( \{ \beta \} \) has the period 1, and as \((a, q) = 1 \) \( m + ax \) runs through a complete residue system mod \( q \), so we can write
\[
\{ y + ax \} = \{ \frac{y(x) + \theta(x)}{q} \}
\]
where \( y(x) \) runs in a certain order over all integers \( \not\equiv 0, 1, \ldots, q-1 \).

If \( 3 \leq y(x) \leq q-3 \), that is always with five exceptions, one can write
\[
\frac{y(x)}{2q} \leq \frac{y(x)}{q} < y(x) + \theta(x) < y(x) + \frac{q}{2} \leq q - \frac{q - y(x)}{2}
\]
Further
\[
\frac{y(x)}{2q} < \frac{y(x) + \theta(x)}{q} < 1 - \frac{q - y(x)}{2q}
\]
and hence
\[
\{ y + ax \} > \min \left( \frac{y(x)}{2q}, \frac{q - y(x)}{2q} \right)
\]
thus
\[
\frac{1}{2q} \max \left( \frac{1}{y(x)}, \frac{1}{q - y(x)} \right)
\]
Therefore (recalling the five exceptions) we have
\[
\sum_{x=0}^{q-1} \min \left( q, \frac{1}{y + ax} \right) \leq 5q + 2q \sum_{y=3}^{q-3} \left( \frac{1}{y} + \frac{1}{q-y} \right) < 5q \frac{\log q}{\log 6} + 4q \frac{q}{q-2} \frac{\log q}{\log 6}
\]
Now we use
\[
\frac{1}{y} \leq \int_{\frac{1}{y}}^{\frac{1}{y+1}} \frac{d\xi}{\xi} = \log y
\]
This leads to
\[
\sum_{x=0}^{q-1} \min \left( q, \frac{1}{y + ax} \right) < \frac{5q \log q}{\log 6} + 4q \log q
\]
\[
< 8q \log q.
\]
Theorem 9.2. If \( g \geq q > 1, (a, q) = 1, \) real, \( |a - \frac{a}{q}| \leq \frac{1}{q^2}, \) then

\[
\sum_{z} \min(q, \frac{1}{q^2 x^2}) < 16g \log q,
\]
where \( z \) runs over \( g \) consecutive integers.

Proof. We divide the numbers \( z \) into groups of \( q \) consecutive numbers, thus we will have \( \left\lfloor \frac{g}{q} \right\rfloor \) groups and at most \( q - 1 \) remaining numbers \( z \). The whole sum thus can be decomposed into \( \left\lfloor \frac{g}{q} \right\rfloor \) sums of the type discussed in the preceding theorem \( (y = \alpha n), \) plus one sum, which is smaller than the sums discussed there. Making use of this decomposition and of Theorem 9.1., we have

\[
\sum_{z} \min(q, \frac{1}{q^2 x^2}) < (\frac{g}{q} + 1) \cdot 8q \log q = 8(g+q) \cdot \log q \leq 16g \cdot \log q.
\]

Theorem 9.3. (Lemma G.) Let \( X > q > 1, (a, q) = 1, \) \( \alpha \) real, \( |a - \frac{a}{q}| \leq \frac{1}{q^2}, \) then

\[
\left| \sum_{v=0}^{q} e^{2 \pi i v x} \right|^2 \leq 32HzX \cdot \log q,
\]
where \( r \) runs over \( h \) different integers \( \leq q \) and \( v \) runs over \( Z \) different integers \( \leq X \).

Proof. For real \( \mu \) we always have

\[
\gamma \leq \sum_{x=1}^{q} e^{2 \pi i x / \mu} \leq q
\]
and if \( \mu \) is not an integer

\[
\left| \sum_{x=1}^{q} e^{2 \pi i x / \mu} \right| = \left| \frac{e^{2 \pi i \mu} - e^{2 \pi i (q+1) \mu}}{1 - e^{2 \pi i \mu}} \right| \leq \frac{2}{|2 \sin(\pi \mu)(\sin(\pi \mu) - i \cos(\pi \mu))|} \leq \frac{1}{\csc(\pi \mu)}.
\]
The function $| \sin nx |$ as a function of $x$ has the period 1, and is an even function, hence
\[
\frac{1}{| \sin nx |} = \frac{1}{| \sin \pi (n - [n]) |} = \frac{1}{| \sin \pi (n - [n]) |}
\]
\[
= \frac{1}{\sin \pi (\lfloor n \rfloor + 1 - \lfloor n \rfloor)} = \frac{1}{\sin \pi [n]}
\]
Now it is well known, that for $0 \leq x \leq \frac{\pi}{2}$, $\sin x < \frac{x}{\pi}$; hence
\[
\frac{1}{\sin \pi [n]} < \frac{2}{\pi [n]^2} < \frac{2 [n]}{\pi [n]^2} \leq \frac{1}{[n]}
\]
(0 ≤ [n] ≤ 1)
So we finally get for $\mu$ not an integer
\[
\left| \sum_{x=1}^{q} e^{2\pi i x \mu} \right| < \frac{1}{[\mu]^2}
\]
Now
\[
\left| \sum_{v} e^{2\pi i v \alpha} \right|^2 = \left| \sum_{v} e^{2\pi i v \alpha} \right|^2 \leq \left( \sum_{v} \left| e^{2\pi i v \alpha} \right| \right)^2
\]
To this sum we apply Schwarz's Inequality
(i.e. $\left( \sum_{v} a_v b_v \right)^2 \leq \left( \sum a_v^2 \right) \left( \sum b_v^2 \right)$), where we put
$a_v = 1$, $b_v = \left| \sum_{v} e^{2\pi i v \alpha} \right|$. Thus we get
\[
\left| \sum_{v} e^{2\pi i v \alpha} \right|^2 \leq H \sum_{v} \left| e^{2\pi i v \alpha} \right|^2 \leq H \sum_{x=1}^{q} \left| e^{2\pi i x \alpha} \right|^2
\]
\[
= H \sum_{x=1}^{q} \sum_{v} e^{2\pi i x (v - v')} \leq H \sum_{v} \min_{x=1}^{q} \left| e^{2\pi i x (v - v')} \right|^2
\]
If we set $v - v' = z$, then
\[
-[x] + 1 \leq z \leq [x] - 1
\]
Every one of these numbers $z$ is obtained at most 2 times from the expression $v - v'$. As
\[
g = 2[x] - 1 \geq 2q - 1 > q
\]
Theorem 9.2 can be applied and yields
\[
\left| \sum_{v} e^{2\pi i v \alpha} \right|^2 \leq H \sum_{z=[x]+1}^{[x]-1} \min_{z=-[x]+1}^{[x]-1} \left( q \right) \leq 32 Hz \log q,
\]
the theorem is thus proved.
Section 10.

**Lemma D (Heilbronn).**

From now on we will set

\[ v = v(k) = \left[ k \log(6k^2) \right] + 1. \]

**Theorem 10.1.**

\[ (1 - \frac{1}{k}) < \frac{1}{6k^2}. \]

**Proof.** Using the power series expansion of \( \log(1 - \frac{1}{k}) \) we have

\[ \log((1 - \frac{1}{k})^v) = v \log(1 - \frac{1}{k}) < -\frac{v}{k}. \]

the definition of \( v \) then gives

\[ \log((1 - \frac{1}{k})^v) < - \log(6k^2) = \log\frac{1}{6k^2}. \]

**Definition 10.1.** For \( h \geq 1 \) we denote by \( \mathcal{A}_h \) the set of all positive numbers \( z \) of the form

\[ z = \sum_{m \geq 0} x_m k^m, \quad x_m \geq 0. \]

**Definition 10.2.** For \( h \geq 1 \) we denote by \( H_h(\psi) \) the number of \( z \leq \psi, z \in \mathcal{A}_h \).

**Theorem 10.2.** For \( h \geq 1, \psi \geq 1, \) we have

\[ H_h(\psi) > \frac{1}{c_{23}(k,h)} \psi^{1-(1 - \frac{1}{k})^h}. \]

**Proof.** The theorem will be proved by induction.

For \( h=1 \) we have

\[ H_1(\psi) = \left\lfloor \frac{1}{k} \right\rfloor \psi^{1-(1 - \frac{1}{k})} = \frac{1}{2} \psi^{1-(1 - \frac{1}{k})}. \]

Now let \( h>1 \) and assume the theorem is true for \( h=1 \), we prove it for \( h \) by considering two cases.
1) \( 1 \leq \psi \leq 8^k \), as \( 1 \leq \phi \), we have
\[
\begin{align*}
H_\phi(\psi) &\geq \frac{1}{c_{25}(k)} \cdot \frac{1}{1 - (1 - \frac{1}{k})^h} \\
&\geq \frac{1}{c_{25}(k)} \cdot \frac{1}{1 - (1 - \frac{1}{k})^h} \\
&\geq \frac{1}{c_{25}(k, h)} \cdot \frac{1}{1 - (1 - \frac{1}{k})^h}.
\end{align*}
\]

2) \( \psi \leq 8^k \), the numbers \( j^k + z \), \( \frac{1}{2} \leq \psi \leq \frac{1}{2} \), \( z \leq (\frac{1}{2} \psi)k - 1 \), \( z \in \mathbb{Z}_k \) are elements of \( \mathbb{Q}_h \), they also are \( \leq \psi \) and different, this is seen as follows:
\[
j^k \leq j^k + z \leq j^k + j^k - 1 < (j+1)k \leq \psi.
\]

Hence we have
\[
H_\psi(\psi) \geq (\frac{1}{2} \psi - 1) - (\frac{1}{2} \psi + 1) \frac{1}{c_{23}(k, h-1)} \cdot \frac{1 - \frac{1}{k}}{1 - (1 - \frac{1}{k})^{h-1}}.
\]
\[
\geq \frac{1}{c_{23}(k, h)} \cdot \frac{1}{1 - (1 - \frac{1}{k})^h}.
\]

**Theorem 10.3. (Lemma D.)** For \( \psi \geq 1 \),
\[
H_\psi(\psi) > \frac{1}{c_{27}(k)} \cdot \frac{1 - \frac{1}{6k^2}}{\frac{1}{6k^2}}.
\]

**Proof.** We substitute \( h = \psi(k) \) in the preceding theorem, and note, that by Theorem 10.1
\[
(1 - \frac{1}{6k^2}) < \frac{1}{6k^2}.
\]

**Section 11.**

**Lemma E (Vinogradov).**

**Theorem 11.1. (Lemma E.)** Let \( \tau \leq \psi \) and let
\[
\begin{align*}
0 &\leq F(\tau) \leq B_j \\
0 &\leq F'(\tau) \leq C \\
0 &\leq F''(\tau),
\end{align*}
\]
Then

$$\left| \sum_{\tau \leq t \leq \tau_2} e^{iF(t)} - \int_{\tau}^{\tau_2} e^{iF(t)} \, dt \right| < 4(1+\sigma+\beta C).$$

Proof. We distinguish three cases.

1) \(\tau_1\) and \(\tau_2\) are integers. Then

$$\left| \sum_{\tau \leq t \leq \tau_2} e^{iF(t)} - \int_{\tau}^{\tau_2} e^{iF(t)} \, dt \right| \leq 1 + \left| \sum_{\tau \leq t \leq \tau_2} e^{iF(t)} \right| - \frac{1}{2} e^{i\overline{F}(\tau_1)} - \frac{1}{2} e^{i\overline{F}(\tau_2)} \left| \int_{\tau}^{\tau_2} e^{iF(t)} \, dt \right|$$

(11.1)

This last result is verified as follows

$$-\int_{\tau}^{\tau_2} \frac{1}{2} \alpha \left( e^{iF(t)} \right) = -\frac{1}{2} e^{i\overline{F}(\tau_2)} + \frac{1}{2} e^{i\overline{F}(\tau_1)}$$

$$\int_{\tau}^{\tau_2} \tau \alpha \left( e^{iF(t)} \right) = \tau_2 e^{iF(\tau_2)} - \tau_1 e^{iF(\tau_1)} - \int_{\tau}^{\tau_2} e^{iF(t)} \, dt)$$

$$-\int_{\tau}^{\tau_2} \frac{1}{2} \alpha \left( e^{iF(t)} \right) = -\sum_{\tau \leq t \leq \tau_2} \frac{1}{2} \alpha \left( e^{iF(t)} \right) = \sum_{\tau \leq t \leq \tau_2} \alpha \left( e^{iF(t)} \right) = \sum_{\tau \leq t \leq \tau_2} (\tau - \tau_1) e^{iF(\tau)} - \tau_2 e^{iF(\tau_1)}$$

Adding these three expressions we see that (11.1) is correct. We next show

$$\left| \int_{\tau}^{\tau_2} (\tau - \tau_2 - \frac{1}{2}) \alpha \left( e^{iF(t)} \right) \, dt \right| \leq \frac{1}{2}$$

to this end we graph the function

the correctness of the bound is then evident.

We now set

$$\varphi(\xi) = \sum_{\tau \leq \xi \leq \tau_2} \left( \tau - \tau_2 - \frac{1}{2} \right) \alpha(\tau) \, dt, \quad \tau \leq \xi \leq \tau_2.$$
From the second mean value theorem follows
\[ |\psi(\xi)| < \frac{1}{\theta} C, \quad \tau_1 \leq \xi \leq \tau_2 \]

The expression \( (11.1) \) then becomes through integration
by parts
\[ = 1 + |\psi(\tau_2) e^{iF(\tau_2)} - 0 - \int_{\tau_1}^{\tau_2} \psi(\tau) iF'(\tau) e^{iF(\tau)} d\tau| \]
\[ \leq 1 + \frac{1}{2} C + \int_{\tau_1}^{\tau_2} \frac{1}{\theta} C F'(\tau) d\tau \leq 1 + \frac{1}{2} C (1 + B) < 4(1 + C + BC) \]

2) \([\tau_1] = [\tau_2]\), then \(\tau_2 - \tau_1 < 1\) and there is no integer \(t\) such
that \(\tau_1 < 4 \leq \tau_2\), hence
\[ |\sum_{\tau_i < \Delta \leq \tau_1} e^{iF(\Delta)} - \int_{\tau_1}^{\tau_2} e^{iF(\tau)} d\tau| = |\int_{\tau_1}^{\tau_2} e^{iF(\tau)} d\tau| < 1 < 4(1 + C + BC) \]

3) \([\tau_1] < [\tau_2]\), then
\[ |\sum_{\tau_i < \Delta \leq \tau_1} e^{iF(\Delta)} - \int_{\tau_1}^{\tau_2} e^{iF(\tau)} d\tau| < \left| \sum_{\tau_i < \Delta \leq \tau_1} e^{iF(\Delta)} - \sum_{[\tau_1]+1}^{[\tau_2]} e^{iF(\tau)} d\tau \right| + 3 \]
and hence by 1)
\[ < 3 + 1 + \frac{1}{\theta} C (1 + B) \leq 4(1 + C + BC). \]
Section 12.

VINOGRADOW'S THEOREM.

The proof we present here is a simplification of Vinogradow's original proof, and is due to Heilbronn. The result we get is not quite as good as Vinogradow's result. He obtained

\[ G(k) \leq 6k \log k + 3k \log 6 + 4k - 2, \]

while we only prove

\[ G(k) \leq 6k \log k + 3k \log 6 + 4k + 3. \]

We require some more notation. We denote by \( N \) an integer \( N \geq 6 \). We further set \( P = Nk \). We then have

\[ \frac{1}{4^2} k - \frac{5}{2} = \frac{1}{4^2} k - \frac{1}{2k} \geq \frac{1}{4^2} Nk \geq \frac{1}{4} > 1. \]

We let

\[ \frac{N}{4} \leq n < N. \]

Now let \( u \) be any element of the set \( \mathcal{G}_u \), then we set for real \( \alpha \)

\[ T(\alpha) = T(\alpha, k, N) = \sum_{m \leq P} e^{2\pi i m^2 \alpha}, \]

\[ R(\alpha) = R(\alpha, k, N) = \sum_{\frac{1}{2} \leq n \leq P} e^{\pi i n^2 \alpha}, \]

\[ S(\alpha) = S(\alpha, k, N) = \sum_{\frac{1}{2} \leq n \leq P} \sum_{1 \leq m \leq P} e^{2\pi i m^2 \alpha} \]

and for real

\[ U(\beta) = U(\alpha, k, N) = \int e^{2\pi i \alpha k \beta} d\omega, \]

\[ V(\beta) = V(\alpha, k, N) = \frac{1}{N} \sum_{1 \leq m \leq P} e^{2\pi i m\beta} \]

Theorem 12.1. The integral

\[ \int_0^1 T(\alpha) R^2(\alpha) S(\alpha) e^{-2\pi i N\alpha} d\alpha \]

is equal to the number of solutions of the diophantine equation
\[ x_1^k + \ldots + x_s^k + u + u' + y^k u'' = N, \]

where \( x_j \leq P, u, u', u'' \) are elements of \( \mathbb{Q}^n \), \( u \leq \frac{1}{4} p^k \), \( u' \leq \frac{1}{4} p^k \), \( u'' \leq \frac{1}{2} p^k \) and \( y \leq p^k \).

**Proof.** The integrand is the sum of a number of terms of the form \( e^{\alpha \pi m} \). Now
\[
\int_0^1 e^{\alpha \pi m} \, d\alpha = \begin{cases} 
1 & \text{for } m = 0 \\
0 & \text{otherwise.} 
\end{cases}
\]

Hence the integral (12.1) is equal to the described number.

Those numbers \( N \) for which the integral (12.1) is different from zero, can therefore be decomposed into the sum of \( s + 3v \) \( k \) th powers of this restricted nature.

Waring's Problem then is solved with
\[ G(k) \leq s + 3v, \]
if it can be shown, that for all \( N \geq N_k \) the integral (12.1) \( \neq 0 \).

For if an \( N \) allows a decomposition of this special type into \( s + 3v \) \( k \) th powers, a fortiori, allows some decomposition into \( s + 3v \) \( k \) th powers.

The remainder of this chapter will be needed to show that the integral (12.1) \( \neq 0 \). It is for the evaluation of this integral, that all the apparatus, we have already built up, is needed.
We denote by $R'$ the number of terms of the sum $R(\alpha)$ and by $S'$ the number of terms of the sum $S(\alpha)$. We have evidently

$$R' = R_v(\frac{1}{4}P^k),$$
$$S' = \left[ \frac{1}{2(k+1)} \right] H_v(\frac{1}{4}P^k - \frac{1}{2}).$$

**Theorem 12.2.**

$$R' \sqrt{S'} > \frac{1}{628(k)} P^{2k} - \frac{1}{4} \log P.$$ 

**Proof.** We use Lemma $D$. (If $P^k > \frac{1}{2} P^{2k}$).

$$R' \sqrt{S'} > \frac{1}{C_{27}(k)} \left( \frac{1}{4} P^{2k} \right)^{1 - \frac{1}{2} P^{-1}} \sqrt{\frac{1}{2} P^{2k} \left( \frac{1}{4} P^{2k-1} \right)^{1 - \frac{1}{2} P^{-1}}}$$

$$> \frac{1}{C_{27}(k)} P^{2k} - \frac{1}{4} P^{-1} + \frac{1}{4} P^{-1} - \frac{1}{4} - \frac{1}{12} P^{-1} + \frac{1}{24} P^{-2}$$

$$= \frac{1}{C_{27}(k)} P^{\frac{3}{2} P^{-1} + \frac{1}{24} P^{-2}} > \frac{1}{C_{28}(k)} P^{\frac{3}{2} P^{-1} + \frac{1}{4} \log P}. \quad (P^{\frac{3}{2} P^{-1} + \frac{1}{4} \log P} > C(k))$$

We now divide the interval of integration of the integral (12.1) $0 \leq \alpha \leq 1$, into subintervals as follows: Let $\frac{a}{q}$ be the elements of The Farey Series belonging to $[P^k - \frac{1}{2}]$, their medians then divide the line segment into the desired subintervals. The first median lies between the fractions $\frac{0}{q}$ and $\frac{1}{q}$, hence it is $\frac{1}{0 + q}$, we prefer to use the interval $[\frac{1}{q}, \frac{1}{q+1}]$ as interval of integration, this is possible as the integrand of (12.1) has the period 1.
For all $\frac{9}{q} \leq \alpha \leq \frac{1}{q}$ we then have
\[ \alpha = \frac{a}{q} + \beta \]
where $1 \leq a \leq q \leq p$ and $\frac{k}{q} - \frac{1}{2} \leq \beta \leq \frac{1}{2}$

where $\frac{a}{q} - \frac{1}{q}$ and $\frac{a+q}{q}$ are the mediants to the left and right of $\frac{a}{q}$. Hence by Theorem 6.1 and Theorem 6.2 we have
\[ \frac{1}{q} \leq \beta \leq \frac{1}{q} \]

A fortiori
\[ |\beta| = \left| \alpha - \frac{a}{q} \right| \leq \frac{1}{q} \leq \frac{1}{q} \]

Those subintervals for which $1 \leq q < \frac{1}{2}$ we call major arcs and the remaining ones minor arcs.

Theorem 12.3. For $\beta$ as defined above
\[ |U(\beta) - V(\beta)| < 4. \]

Proof. A change of variable $\lambda = \frac{e^{2\pi i \beta}}{\lambda}$ gives
\[ |U(\beta) - V(\beta)| = \left| \int_{0}^{N} e^{2\pi i \beta \lambda^{2}} \lambda^{2} d\lambda - \left( \sum_{k=1}^{N} \frac{1}{\lambda^{k}} \right) e^{2\pi i \beta \lambda^{2}} \lambda^{2} d\lambda \right| \]
\[ \leq \left| \int_{0}^{N} e^{2\pi i \beta \lambda^{2}} \lambda^{2} d\lambda - \left( \sum_{k=1}^{N} \frac{1}{\lambda^{k}} \right) e^{2\pi i \beta \lambda^{2}} \lambda^{2} d\lambda \right| \]
\[ + \frac{1}{\lambda^{k}} \int_{0}^{\lambda^{2}} \lambda^{2} d\lambda + \frac{1}{\lambda^{k}} \int_{1}^{N} \lambda^{2} d\lambda \]
\[ \leq 1 + \frac{1}{\lambda^{k}} + \frac{1}{\lambda^{k}} \left| \int_{0}^{N} (e^{2\pi i \beta \lambda^{2}} - 1) d\lambda \right| \]
then as $|\lambda - [\lambda] - \frac{1}{2}| \leq \frac{1}{2}$ we have
\[ \leq \frac{4}{3} + \frac{1}{2k} \int_{0}^{N} \lambda^{2} d\lambda + \frac{1}{2k} \int_{1}^{N} \lambda^{2} d\lambda \]
\[ < \frac{4}{3} + \frac{7|\beta|}{2k} \int_{0}^{N} \lambda^{2} d\lambda + \frac{1}{2k} \int_{1}^{N} \lambda^{2} d\lambda \]
\[
\frac{1}{2} + \frac{7}{2} |\beta| N^\frac{1}{2} + \frac{1}{2} \frac{3}{2} \leq \frac{1}{2} + \frac{7}{2} \frac{N^\frac{1}{2}}{P^\frac{1}{2}} + \frac{1}{2}
\]
\[
\frac{1}{12} + \frac{7}{2} \frac{1}{N^\frac{1}{2} \sqrt{c}} \leq \frac{1}{12} + \frac{7}{2} \frac{1}{2^{1/3}} < 4.
\]

**Theorem 12.4.** For \(\alpha, \beta\) and \(\frac{a}{q}\) as defined above,

and for

\[
\mathcal{S} = e^{2\pi i \frac{a}{q}},
\]

we have

\[
|T(\alpha) - q^{-1} S_p U(\beta)| < c_{30}(k) q.
\]

**Proof.** We decompose \(T(\alpha)\) as follows

\[
T(\alpha) = \sum_{q=1}^{Q} \sum_{-\frac{a}{q} < \lambda \leq \frac{P-r}{q}} e^{2\pi i (\lambda q + \tau) k} = \sum_{q=1}^{Q} \sum_{-\frac{a}{q} < \lambda \leq \frac{P-r}{q}} e^{2\pi i (\lambda q + \tau) k}
\]

The transformation \(x = \frac{a}{q} \cdot x\) gives

\[
q^{-1} U(\beta) = \frac{1}{q} \int_0^P e^{2\pi i \omega \cdot \beta} d\omega = \frac{1}{q} \int_0^P e^{2\pi i (\omega q + \tau) k \overline{\beta}} d\omega.
\]

Thus we obtain

\[
T(\alpha) - q^{-1} S_p U(\beta)
\]

\[
= \sum_{q=1}^{Q} \sum_{-\frac{a}{q} < \lambda \leq \frac{P-r}{q}} e^{2\pi i (\lambda q + \tau) k} - \frac{1}{q} \int_0^P e^{2\pi i (\omega q + \tau) k \overline{\beta}} d\omega.
\]

To this difference we can apply Lemma E with

\[
t_1 = \frac{a}{q}, \quad t_2 = \frac{P-r}{q},
\]

\[
F(t) = \omega^t (\omega q + \tau) k \overline{\beta}.
\]

For we have for \(t_1 \leq t \leq t_2\)

\[
0 \leq F(t) \leq 2\omega \frac{P-k}{k} |\beta|,
\]

and as \(|\beta| < \frac{1}{qP^{-\frac{1}{2}}}\) we have

\[
0 \leq F(t) \leq 2\omega \frac{P\frac{1}{2}}{q} \leq 2\omega \frac{P\frac{1}{2}}{q},
\]

and

\[
0 \leq F'(t) \leq 2\omega k q P^{k-1} |\beta| \leq 2\omega k P^{-\frac{1}{2}}
\]

and finally

\[
0 \leq F''(t).
\]
We now note that we may replace $\beta$ by $|\beta|$, for we have $|\Sigma e^\alpha| = |\Sigma e^\beta|$. From Lemma $E$ we therefore obtain

$$|T(\alpha) - q^{-1}\Sigma e^\beta| \leq \frac{q}{2} \left| \sum_{t_1 \leq t_2} e^{iF(t_1)} - \int_{t_1}^{t_2} e^{iF(t)} dt \right|$$

$$\leq 4q \left( t + 2\pi k \right)^{\frac{1}{2}} + 4\pi k < 4q \left( t + 2\pi k + 4\pi^2 k \right) < c_{30}(k) q.$$  

**Theorem 12.5.**

$$|T(\alpha) - q^{-1}\Sigma e^\beta| < c_{31} \cdot q.$$  

**Proof.** Evidently $|\Sigma e^\beta| < q$, hence

$$|T(\alpha) - q^{-1}\Sigma e^\beta| \leq |T(\alpha) - q^{-1}\Sigma e^\beta| + |q^{-1}\Sigma e^\beta| |U(\beta) - V(\beta)|$$

$$\leq c_{30} \cdot q + 4 < c_{31} \cdot q.$$  

**Theorem 12.6.** On every $\Omega$, 

$$\max(|T(\alpha)|, |q^{-1}\Sigma e^\beta|) < c_{32}(k) q \cdot \frac{1}{k} p.$$  

**Proof.** Let 

$$W(\alpha) = T(\alpha) - q^{-1}\Sigma e^\beta,$$

then it follows from the preceding theorem ($\alpha$ on $\Omega$, hence $q \leq \frac{1}{k}$) that

$$|W(\alpha)| < c_{31} \cdot q \leq c_{31} \cdot q \cdot \frac{1}{k} p \cdot \frac{q^2}{P}$$

$$< c_{31} \cdot q \cdot \frac{1}{k} p.$$  

For $V(\beta)$ we have evidently

$$|V(\beta)| \leq \frac{1}{k} \sum_{m=1}^{N} \sum_{m=1}^{N} m^{k-1} \cdot \frac{1}{k} \int_{0}^{N} m^{k-1} \cdot \gamma = P$$

Hence using Lemma $A$, we get

$$|q^{-1}\Sigma e^\beta| < c_{4}(k) q \cdot \frac{1}{k} p.$$  

We then have

$$|T(\alpha)| \leq |q^{-1}\Sigma e^\beta| + |W(\alpha)| < c_{31} \cdot q \cdot \frac{1}{k} p + c_{31} \cdot q \cdot \frac{1}{k} p.$$  

The theorem then follows immediately.

**Theorem 12.6.** On every $\mathcal{A}$t,

$$|T^s(\alpha) - q^{-s}s P V^s(\beta)| \leq c_{33}(k) q^{-2} p^{4k-1}.$$  

**Proof.** Using the two preceding theorems

we have

$$|T^s(\alpha) - q^{-s}s P V^s(\beta)|$$

$$= \left| w(\alpha) \sum_{n \in \mathcal{A}} T^{s-1-m}(\alpha) (q^{-1}s V(\beta))^{m} \right|$$

$$\leq c_{31}(k) q s (\max(\|T(\alpha)\|, |q^{-1}s P V(\beta)|))^{s-1}$$

$$\leq c_{33}(k) q^{s} (q^{-1} + 1) p^{s-1}$$

$$= c_{33}(k) q^{-3/2} + \frac{1}{k} p^{4k-1}$$

$$< c_{33}(k) q^{-2} p^{4k-1}.$$  

**Definition 12.1.**

$$D(k,n) = \frac{1}{k} \sum_{m_{j}} \left( \frac{s}{m_{j}} \right)^{\frac{1}{k} - 1},$$

where the summation extends over all sets $m_{j}$ with $\sum_{j=1}^{\infty} m_{j} = n$, $m_{j} > 0$.

**Theorem 12.7.**

$$D(k,n) < c_{34}(k) p^{3k}.$$  

**Proof.** There are less than $n^{3-1}$ terms in

$D(k,n)$, for every one there is a term in the sum

$$(12.2) \frac{1}{k} \sum_{m_{i}, \ldots, m_{j}} \left( \frac{s}{m_{i}} \right)^{\frac{1}{k} - 1} (\frac{n}{s})^{\frac{1}{k} - 1}$$

that is greater than the corresponding term in $D(k,n)$.

For we have in $\sum_{j=1}^{\infty} m_{j} = n$ at least one $m_{j}$ with $m_{j} \geq \frac{n}{s}$, hence

$$(m_{j})^{\frac{1}{k} - 1} \leq (\frac{n}{s})^{\frac{1}{k} - 1}.$$
We can arrange the comparison in such a way, that any two terms of $D(k,n)$ will be compared with different terms of the sum (12.2) and there still will be positive terms left in (12.2). Hence

$$D(k,n) \leq \frac{1}{k} s \sum_{m_j} \left( \frac{s}{2} m_j \right)^{1-k} \left( \frac{n}{s} \right)^{s-1}$$

$$\leq c_{35}(k) \left( \sum_{m_j} \frac{s}{2} m_j \right)^{s-1}$$

using the approximation

$$\sum_{m_j} \frac{s}{2} m_j \leq \int \frac{n}{s} \frac{1}{2} m_j^s \, \mathrm{d}m_j$$

we then have

$$D(k,n) \leq c_{35}(k) n^{k-1} \left( \frac{nk}{s} \right)^{s-1}$$

$$= c_{34}(k) n^{k-1}$$

$$= c_{34}(k) n^3 < c_{34}(k) p^{3k}$$

**Theorem 12.8.** For $n \geq s$

$$D(k,n) > \frac{1}{c_{36}(k)} p^{3k}.$$  

**Proof.** We have

$$\left( \frac{s}{2} m_j \right) \leq \left( \frac{s}{2} \sum_{j=1}^n m_j \right)^{s-1}$$

that is the arithmetic mean is greater than or equal to the geometric mean, hence $(\frac{1}{k} - 1 < 0)$

$$D(k,n) \geq \frac{1}{k} s \sum_{m} \left( \left( \frac{s}{2} \sum_{j=1}^n m_j \right)^{s-1} \right)^{k-1}$$

where the sum on the right is over all terms with $m_j > 0$, $\sum_{j=1}^n m_j = n$. We therefore have

$$D(k,n) > \frac{1}{c_{37}(k)} \left( \frac{n}{s} \right)^{s-1} \left( \frac{1}{2} \right)^{s-1}.$$  

The last sum is for $n \geq s$ over at least $\left[ \frac{n}{s} \right]^{s-1} > (\frac{1}{2} s) = \frac{n}{s} = c(k)n^{s-1}$ terms, hence

$$D(k,n) > \frac{1}{c_{38}(k)} n^{4-s} n^{s-1} = \frac{n^3}{c_{38}(k)}$$

$$\geq \frac{1}{c_{38}(k)} \left( \frac{p^k}{4} \right)^3 = \frac{1}{c_{36}(k)} p^{3k}.$$
Theorem 12.9. For $0 < |\beta| \leq \frac{1}{2}$,

$$|V(\beta)| < 2 |\beta|^{-\frac{1}{2}}$$

Proof. For $a \leq b$ we have

$$|\sum_{m=a}^{b} e^{2\pi i \beta m}| = |\sum_{x=1}^{b-a+1} e^{2\pi i \beta (x+a-1)}|$$

$$= |\sum_{x=1}^{b-a+1} e^{2\pi i \beta x}| < \frac{1}{|\beta|} = \frac{1}{|\beta|}$$

this was shown in the proof of Lemma C. If we set

$$M = \min(\frac{1}{|\beta|}, \bar{N})$$

we can write

$$|V(\beta)| = \left| \frac{1}{\beta} \sum_{m=1}^{M} m^{-1} e^{2\pi i \beta m} \right| \leq \frac{1}{\beta} \sum_{m=M}^{\bar{N}} m^{-1} + \sum_{m=M}^{\bar{N}} m^{-1} e^{2\pi i \beta m}$$

For the first of these sums we use the integral approximation

$$\frac{1}{\beta} \sum_{m=M}^{\bar{N}} m^{-1} < \frac{1}{\beta} \int_{M}^{\bar{N}} x^{-1} dx \leq |\beta|^{-\frac{1}{2}}$$

For the second sum we obtain by the help of the Abel transformation ($a_n = \sum_{m=1}^{M} a_m$, $a_n = e^{2\pi i \beta n}$, $b_n = n^{-1}$, $b_n$)

$$\left| \sum_{n=1}^{N} a_n b_n \right| < \max |A_n| (\sum_{n=1}^{N} (b_n - b_{n+1}) + b_{n+1}) = \max |A_n| \cdot b_{n+1}$$

Thus

$$|V(\beta)| < |\beta|^{-\frac{1}{2}} + \frac{1}{|\beta|} (|\beta|^{-\frac{1}{2}}) = 2 |\beta|^{-\frac{1}{2}}$$

Theorem 12.10. For $0 < |\beta| \leq \frac{1}{2}$,

$$|V^S(\beta)| < c_{39}(k) |\beta|^{-4}.$$ 

Proof. From the preceding theorem we get

$$|V^S(\beta)| < 2^s |\beta|^{-4} = c_{39}(k) |\beta|^{-4}.$$
Theorem 12.11.
\[ \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-2\pi\eta \beta} V^5(\beta) \, d\beta = D(k,n) \]

Proof. The theorem then follows from an argument similar to the one used in the proof of Theorem 12.1, if it is noted that \( n < N \). Hence the integral will give all the terms of the sum \( D(k,n) \).

Theorem 12.12. For every \( \mathcal{M} \),
\[ |\int_{\mathcal{M}} e^{-2\pi\eta \beta} V^5(\beta) \, d\beta - D(k,n)| < c_{40}(k) q P^{3k-\frac{1}{2}} \]

Proof. On every \( \mathcal{M} \), \( |\beta| \leq \frac{1}{P} < \frac{1}{2} \). From the preceding theorem we have then
\[ |\int_{\mathcal{M}} e^{-2\pi\eta \beta} V^5(\beta) \, d\beta - D(k,n)| \leq \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| V^5(\beta) \right| \, d\beta + \int_{\frac{1}{2}}^{1} \left| V^5(\beta) \right| \, d\beta \]

and from Theorem 12.10
\[ \leq 2 \int_{\frac{1}{2}}^{1} c_{34}(k) \beta^{-4} \, d\beta < 2 c_{34}(k) \int_{1/2}^{1} \beta^{-4} \, d\beta < c_{40}(k) q^3 P^{3k-\frac{3}{2}} \]

but we are on an \( \mathcal{M} \), hence \( q < P^2 \), and our expression is
\[ < c_{40}(k) q P^{3k-\frac{1}{2}} \]

Theorem 12.13. For every \( \mathcal{M} \),
\[ |\int_{\mathcal{M}} e^{-2\pi\eta \alpha} T^5(\alpha) \, d\alpha - q^{-5} S^5 P^{-n} D(k,n)| < c_{41}(k) q^{-3} P^{3k-\frac{1}{2}} \]

Proof. The length of any interval \( \mathcal{M} \) is \( \leq \frac{2q}{q^2+1} \).

Now we note
\[ e^{-2\pi\eta \alpha} = e^{-2\pi\eta \beta \alpha} \eta e^{-2\pi\eta \beta n}, e^{-2\pi\eta \beta n} = \delta^\eta e^{-2\pi\eta \beta n} \]
Hence we obtain from Theorem 12.6
\[ |e^{-2\pi i n a T^s(\alpha)} - q^{-s} \sum p^n e^{-2\pi i n \beta V^s(\beta)}| < c_{33}(k) q^{-2} p^{4k-1} \]
and hence

\[ (12.3) \left| \int_{\mathbb{R}^+} e^{-2\pi i n a T^s(\alpha)} d\alpha - \int_{\mathbb{R}^+} q^{-s} \sum p^n e^{-2\pi i n \beta V^s(\beta)} d\beta \right| < c_{42}(k) q^{-3} p^{3k-\frac{1}{2}} \]

Now using Lemma A and Theorem 12.12 we get

\[ (12.4) \left| \int_{\mathbb{R}^+} q^{-s} \sum p^n e^{-2\pi i n \beta V^s(\beta)} d\beta - q^{-s} \sum p^n D(h, n) \right| < c_{40}(k) p^{3k-\frac{1}{2}} q c_{4}(k) (q^{1-\frac{1}{2}}) q^{-3} < c_{4}(k) c_{40}(k) q^{-3} p^{3k-\frac{1}{2}} \]

The combination of the inequalities (12.3) and (12.4)

then gives the desired result

**Theorem 12.14.**
\[ \left| \sum p^n \int_{\mathbb{R}^+} e^{-2\pi i n a T^s(\alpha)} d\alpha - D(h, n) \sum_{q < p^{\frac{1}{2}}} A_q \right| < c_{45}(k) p^{3k-\frac{1}{2}} \]

**Proof.** Recalling the definition of \( a_q \) (Definition 7.2), and noting that \( \sum \) means summation over all \( \frac{a}{q} \), with \( (a, q) = 1 \), \( q < p^{\frac{1}{2}} \), we have as a consequence of the preceding theorem

\[ (12.5) \left| \sum p^n \int_{\mathbb{R}^+} e^{-2\pi i n a T^s(\alpha)} d\alpha - D(h, n) \sum_{q < p^{\frac{1}{2}}} A_q \right| < c_{44}(k) p^{3k-\frac{1}{2}} \]

From Theorem 8.1 and Definition 8.1 we have

\[ \left| \sum_{q < p^{\frac{1}{2}}} A_q - \mathcal{G}(h, n) \right| \leq c_{20}(k) \sum_{q < p^{\frac{1}{2}}} q^{-3} < c_{45}(k) p^{\frac{1}{2}} \]

now using Theorem 12.7 we get

\[ (12.6) \left| D(h, n) \sum_{q < p^{\frac{1}{2}}} A_q - D(h, n) \mathcal{G}(h, n) \right| < c_{46}(k) p^{3k-\frac{1}{2}} \]

The theorem follows from (12.5) and (12.6).
Theorem 12.15. If \( N > c_{47}(k) \), then
\[
\mathcal{R} \sum \int_{\alpha} e^{-2\pi i n \alpha} T^S(\alpha) d\alpha > \frac{1}{C_{48}(k)} P^3 k
\]

Proof. From the preceding theorem follows immediately \((\mathcal{V}(k,n) \text{ real positive})\)
\[
\mathcal{R} \sum \int_{\alpha} e^{-2\pi i n \alpha} T^S(\alpha) d\alpha > D(k,n) \mathcal{V}(k,n) - C_{43}(k) P^{3k-\frac{1}{2}}
\]
and thus from Lemma B and from Theorem 12.8 we have for \( n > s \), that is \( N > s \)
\[
\mathcal{R} \sum \int_{\alpha} e^{-2\pi i n \alpha} T^S(\alpha) d\alpha > \frac{1}{C_{36}(k)} P^{3k} \cdot \frac{1}{C_5(k)} - \frac{1}{C_{43}(k)} P^{3k-\frac{1}{2}}
\]
\[
= \frac{P^{3k}}{C_{44}(k)} (1 - \frac{C_5(k)}{C_5(D)})
\]
hence for \( N > c_{47}(k) > (2c_{50}(k))^2k \), \( (>s) \), the real part of the integral is
\[
> \frac{P^{3k}}{2C_{44}(k)}
\]

Theorem 12.16. For \( N > c_{47}(k) \)
\[
\mathcal{R} \sum \int_{\alpha} e^{-2\pi i N \alpha} T^S(\alpha) R^2(\alpha) S(\alpha) d\alpha > \frac{1}{C_{48}(k)} P^{3k} R^2 S^1
\]

Proof.
\[
e^{-2\pi i N \alpha} T^S(\alpha) R^2(\alpha) S(\alpha) = \sum_{n_1, n_2, n_3} e^{-2\pi i (N - n_1 - n_2 - n_3) \alpha} T^S(\alpha)
\]
this is a sum of \( R^2 S^1 \) terms. We write
\[
n = N - u_1 - u_2 - y^k u_3,
\]
and have (referring back to the Definitions of \( R \) and \( S \) at the beginning of this section)
\[
N n N - \frac{1}{4} N - \frac{1}{4} N - \frac{1}{4} y^k P \cdot \frac{1}{2} = \frac{N}{4}.
\]
Hence Theorem 12.15 can be used for \( N > c_{47}(k) \), this
yields
\[
\mathcal{R} \sum \int_{\alpha} e^{-2\pi i N \alpha} T^S(\alpha) R^2(\alpha) S(\alpha) d\alpha = \sum_{n_1, n_2, n_3} \mathcal{R} \sum \int_{\alpha} e^{-2\pi i n_1 \alpha} T^S(\alpha) d\alpha
\]
\[
> \frac{1}{C_{48}(k)} P^{3k} R^2 S^1.
\]
Theorem 12.17.

\[ \int_0^1 R^2(\alpha) \, d\alpha = R' \]

**Proof.**

\[ \int_0^1 R^2(\alpha) \, d\alpha = \int_0^1 \sum_{m \in \mathbb{Z}^N} e^{2\pi im \alpha} \, d\alpha = \sum_{m \in \mathbb{Z}^N} \int_0^1 e^{2\pi im \alpha} \, d\alpha = R' \quad (u_1-u_2=0 \text{ has } R' \text{ solutions}) \]

Theorem 12.18.

\[ \sum_m \sum_n \left| T^S(\alpha) R^2(\alpha) S(\alpha) \right| \, d\alpha < c_{51}(k) R^{3k} R'^2 S' \frac{1}{\log P} \]

**Proof.** Evidently \(|T(\alpha)| \leq P\). We apply Lemma C with \(r=y^k, v=u, A=\lfloor P^2k \rfloor, Hz=S', x=P^k-\frac{1}{2}, \) here \(r \leq \frac{r}{2} \leq q \leq \frac{P^k-1}{2}, q > 1\). This gives

\[ |S(\alpha)|^2 = \left| \sum_{n \in \mathbb{Z}} e^{2\pi i n \alpha} \right|^2 = 32 \leq \frac{P^k-\frac{1}{2}}{P} \log (P^{\frac{k}{2}}) \]

Successive application of this result, Theorem 12.17 and Theorem 12.2 then yields

\[ \sum m \sum n \left| T^S(\alpha) R^2(\alpha) S(\alpha) \right| \, d\alpha \leq c_{52}(k) P^{4k} P^k \frac{1}{2} \sqrt{S'} \sqrt{\log P} \int_0^1 R^2(\alpha) \, d\alpha \leq c_{52}(k) P^{4k} P^k \frac{1}{2} \sqrt{S'} \sqrt{\log P} \leq c_{52}(k) P^{4k} P^k \frac{1}{2} \sqrt{S'} \sqrt{\log P} \leq c_{51}(k) P^{3k} R'^2 S' \frac{1}{\log P} \]

Theorem 12.19.

\[ \Re \sum_m e^{-2\pi i m \alpha} T^S(\alpha) R^2(\alpha) S(\alpha) \, d\alpha > -c_{51}(k) P^{3k} R^2 S' \frac{1}{\log P} \]

**Proof.** Together with the preceding theorem.
**Theorem 12.20.** For \( N > c_{53}(k) \),

\[
\int_0^1 e^{-2\pi i n x} T^*(a) R^2(a) S(a) \, da \neq 0
\]

**Proof.**

\[
\mathcal{R} f'_0 = \mathcal{R} f'_1 = \mathcal{R} \sum f_m + \mathcal{R} \sum f_m = (f'_0)
\]

hence for \( N > c_{47}(k) \)

\[
\mathcal{R} f'_1 > \frac{1}{c_{48}(k)} p^{3k} R^2 S^1 - c_{51}(k) p^{3k} R^2 S^1 \frac{1}{\sqrt{\log p}}
\]

\[
= \frac{1}{c_{48}(k)} p^{3k} R^2 S^1 \left(1 - \frac{c_{54}(k)}{\sqrt{\log p}} \right)
\]

This expression is certainly greater than zero, if

\[ N > c_{53}(k) = \max \left( c_{47}(k), e^{c_{54}(k)} \right) \]

This constant is very generous. Vinogradov's original proof gave a much smaller value and is therefore preferable as far as the determination of \( g(k) \) is concerned.

**Theorem 12.21.** For \( k \geq 3 \),

\[ G(k) \leq 6k + \log k + 3k \log 6 + 4k + 3. \]

**Proof.** Obvious.
Section 13.

A COMPARISON.

The so-called "asymptotic theory" in connection with Waring's Problem is due to Hardy and Littlewood. This is clear even in the development of Vinogradov's result, which we gave in the preceding pages, to illustrate that we only need to refer to the two basic lemmas of Hardy and Littlewood, which we used. It seems however of interest to name the improvements which enabled Vinogradov to obtain so much better results in so much shorter space. To this end we describe how Hardy and Littlewood approached the problem.

Let \( r_s(n) \) be the number of solutions of the diophantine equation

\[
\sum_{m=1}^{s} x^k = n, \quad x_m > 0.
\]

Hardy and Littlewood considered the following expression for \( r_s(n) \)

\[
(13.1) \quad r_s(n) = \frac{1}{2\pi i} \oint_{\Gamma} \frac{f_s(x)}{x^{n+1}} \, dx
\]

where \( f_s(x) = \sum_{k=1}^{s} x^k \). It is then easily seen that

\[
f_s(x) = \sum_{n=1}^{\infty} r_s(n) \, x^n,
\]

and (13.1) is immediately clear. The radius of convergence of \( f(x) \) is 1, so the circle \( x = 1 \) can not be used as a path of integration. Hardy and Littlewood
used \( x = e^{\frac{1}{n}} \). Thus we get
\[
(13.2) \quad r_s(n) = \sum_{a=1}^{\infty} e\left(-\frac{1}{n} + 2\pi i \alpha\right)^{\lambda} e\left(-\frac{1}{n} + 2\pi i \alpha\right)^{m} \alpha \lambda
\]

In the difference between this expression and the integral (12.1) is contained the difference in the method of evaluation.

Vinogradov's first improvement was a simple one, he considered finite series instead of infinite ones, thus avoided convergence questions and also was able to avoid the real term in the exponent of \( e \) in his integrand. This simplified matters, but would scarcely have given rise to improved results.

The second difference is that the integrand Vinogradov uses is composed of three different types of sums. This of course gives rise to a different decomposition of \( n \), actually as we noted a more special one than the decomposition required for Waring's Problem. The type of integrand and therefore the special decomposition is needed because only to a sum like \( S(\lambda) \) does Vinogradow's Lemma apply. Hardy and Littlewood used in its place a lemma due to Weyl, which gives a bound for sums of the type
\[
\left| \sum_{\lambda=0}^{M} e^{2\pi i \lambda \alpha} \lambda^k \right|
\]
in terms of \( M \) and \( q \). As we already mentioned, Vinogradow's Lemma is so much better than Weyl's
that it is this improvement in method that gave rise to the improved results he obtained.

Section 14.

LOWER BOUNDS FOR $G(k)$ AND $g(k)$.

**Theorem 14.1.** If $k \geq 2$, then

$$g(k) \geq 2^k + \left[ \left( \frac{3}{2} \right)^k \right] - 2.$$

**Proof.** The number $q = \left( \frac{3}{2} \right)^k$ is the greatest integer satisfying the inequality

$$2^q - 1 < 3^k.$$

The number $2^q - 1$ can only be composed of $k$ th powers of $1$ and $2$. Its decomposition into a minimum number of $k$ th powers is

$$2^q - 1 = \sum_{m=1}^{q-1} 2^k + \sum_{m=1}^{q-1} 1^k.$$

This sum has $2^q + q - 2$ terms.

**Theorem 14.2.** If $k \geq 2$, then

$$G(k) \geq k + 1.$$

**Proof.** If $n = \sum_{m=1}^{k} x_m^k$, $x_m \geq 0$, we certainly must have $x_m \leq nk$. We now find an estimate for the number of integers $0 \leq n \leq y$ that can be expressed as the sum of $k$ $k$ th powers. For every $x_m$ we have $\leq \frac{1}{y^k+1}$ choices and as the order of the $x_m$ can be disregarded, we get as an upper bound, for the number of $n$ expressible in the desired form, $\left( \frac{y^k+1}{k+1} \right)^k < \frac{3}{4} y$, $y > y_o(k)$.

Hence an infinity of numbers $n$ cannot be expressed as the sum of $k$ $k$ th powers, and the theorem is proved.
BIBLIOGRAPHY.

1) A. Cauchy, Demonstration d'un Theoreme curieux sur les nombres, Oeuvres vol 6, part 2, pp. 146-8.
   (This paper contains the proof of the main Theorem on Farey Series.)

   (An elaboration of Vinogradow's result for more general decompositions, and an estimate of the number $N_0$, such that for all numbers greater than $N_0$ a decomposition into $G(k)$ $k$th powers is possible.)


   (These two papers together contain the best results so far obtained on $g(k)$.)

   (This gives an account, without proofs, of the recent results on the problem.)

   (An observation, without proof, on a property of the series which henceforth bore his name.)

8) G. H. Hardy and J. E. Littlewood, Some problems of partitio numerorum,
   part 1, Göttinger Nachrichten (1920) pp 33-54.
   part 6, " " vol 23 (1925), pp. 1-37
   (In these four papers the authors develop the asymptotic theory for the determination of upper bounds of $G(k)$.)


    (Contains a simplified proof of Vinogradov's Theorem.)

    (Prove $G(4) \leq 17$.)

    (Proves $g(k) < \infty$ for all $k$.)

13) A. J. Kempner, Über das Waringsche Problem, Diss. Göttingen 1912.
    (Completes proof of Wieferich's Theorem on $g(3)$.)
14) J. Li La Grange, Demonstration d'un theoreme d'arithmetique, 
(Proves $g(2) = 4$.)

15) E. Landau, Vorlesungen über Zahlentheorie, vol 1, 
Leipzig 1927.

16) E., Über einige neuere Fortschritte der additiven 
Zahlentheorie, Cambridge Tracts in Math. no. 35, 
Cambridge 1937.
(Contains the proof of Vinogradov's Theorem that is 
presented here.)

17) Pillai, Jour. Indian Math. Soc. (2) vol. 2 (1936), 
pp. 16-44.
(Contains the best results on $g(k)$, proved independently 
by Dickson.)

18) H. Rademacher, Trends in Research: The Analytic 
pp. 379-401.
(Gives a discussion of Vinogradov's improvements over 
Hardy and Littlewood.)

(Contains his proof of his Theorem.)

20) E. Waring, Meditationes algebraicae, Cambridge 1770, 
pp. 204-5.
(This is where he made his conjecture.)
(Contains the Lemma used by Hardy and Littlewood to manage the minor arcs.)

(Contains the, not complete, proof that $g(3) \leq 9$.)

This is only a very incomplete Bibliography.